TECHNICAL REPORT 2002-007

Ballistic Missile Single Integrated Air Picture (SIAP) Metrics

NOVEMBER 2002

SINGLE INTEGRATED AIR PICTURE (SIAP)
System Engineering
Task Force (SE TF)
1931 Jefferson Davis Highway
Crystal Mall 3, Suite 1142
Arlington, VA 22203

DISTRIBUTION STATEMENT A: Approved for public release; distribution is unlimited

Report Documentation Page

Report Date	Report Type	Dates Covered (from... to)
00 Nov 2002	N/A	-

Title and Subtitle Ballistic Missile Single Integrated Air Picture (SIAP) Metrics	**Contract Number**
	Grant Number
	Program Element Number
Author(s)	**Project Number**
	Task Number
	Work Unit Number
Performing Organization Name(s) and Address(es) Single Integrated Air Picture (SIAP) System Engineering Task Force (SE TF) 1931 Jefferson Davis Highway Crystal Mall 3, Suite 1142 Arlington, VA 22203	**Performing Organization Report Number** TR2002-007
Sponsoring/Monitoring Agency Name(s) and Address(es)	**Sponsor/Monitor's Acronym(s)**
	Sponsor/Monitor's Report Number(s)
Distribution/Availability Statement Approved for public release, distribution unlimited	
Supplementary Notes	
Abstract	
Subject Terms	
Report Classification unclassified	**Classification of this page** unclassified
Classification of Abstract unclassified	**Limitation of Abstract** UU
Number of Pages 96	

TECHNICAL REPORT 2002-007

Ballistic Missile Single Integrated Air Picture (SIAP) Metrics

NOVEMBER 2002

SINGLE INTEGRATED AIR PICTURE (SIAP)

System Engineering

Task Force (SE TF)

1931 Jefferson Davis Highway
Crystal Mall 3, Suite 1142
Arlington, VA 22203

CDR P. M. VOTRUBA, USN
Analysis Branch Head

CAPT J. W. WILSON, USN
Technical Director

Col H. DUTCHYSHYN, USAF
Deputy System Engineer

RADM M. G. MATHIS, USN
System Engineer

FOREWORD

List of Contributors

The following members of the SIAP SE TF Attributes and Metrics Working Group contributed to this technical report, either through their work on the draft, or through their participation in the review process:

Jeff Heckel	SIAP SE TF (chair)
Gerald Driggers	USA/SMDC
Karlene Duncan	JNIC
Tom Hart	USMC/MITRE
Robert Isbell	MDA/Sparta
John Jordan	USA/CAS
Matt Kelly	MDA/Boeing
Ernie MacNamara	CNA
Richard Nisley	JTAMDO/SAIC
John Nordmann	USAF/MITRE
Crystle Pishon	MITRE
Carl Raiche	JNIC
Ron Rothrock	SIAP SE TF/Sparta
Mike Sapp	USN/APL
Erik Van Fleet	USN/NSWC
Brett Zombro	SIAP SE TF/SPA

EXECUTIVE SUMMARY

PROBLEM

The SIAP System Engineer (SIAP SE) defined a comprehensive set of metrics for quantifying SIAP capability, focused exclusively on the air vehicle component of the SIAP (SIAP SE TF Technical Reports 2001-001, -002, and -003: DTIC References ADA 397215, ADA 397221, and ADA 397225). As planned SIAP assessments also include a ballistic missile component, there is a need to extend the SIAP metrics to cover scenarios in which ballistic missiles/objects as well as air vehicles are present. The extended set of metrics, referred to in this report as "ballistic missile SIAP metrics," should ultimately span the attribute and MOP levels in the SIAP metrics hierarchy defined in SIAP SE TF Technical Report 2001-002.

OBJECTIVES/PURPOSE

The development of ballistic missile SIAP metrics must meet a number of objectives. In the near term, ballistic missile SIAP attributes and corresponding quantitative measures must be defined so that (1) current Joint Requirements Oversight Council (JROC)-validated requirements for theater ballistic missile (TBM) early warning (EW) are fully supportable through assessments based on SIAP attributes, and (2) sufficient metrics guidance is provided to meet the SIAP SE's assessment goals in near-term (Calendar Year 2002) test events involving a ballistic missile component. In addition, (3) ballistic missile SIAP MOPs should be defined so that root-cause and critical experiment analyses for these same events are adequately supported. Over the longer term (one year or more): (4) the ballistic missile SIAP metrics should support thorough end-to-end analyses of SIAP capability, relating engineering improvements to warfighting benefits in a BMD context; (5) the development of SIAP metrics generally should be institutionalized within the Department of Defense and in particular within the joint missile defense communities so that the further evolution of these metrics can be streamlined with the evolution of joint requirements. This report intends to meet the first three (near-term) objectives, and to serve as a contribution to the longer-term goals.

APPROACH

1. Leverage the metrics package used by MDA in the Ballistic Missile Defense Benchmark model.
2. Using current JROC-validated requirements, focusing on SIAP attributes and Ballistic Missile EW requirements in the 2001 Theater Air and Missile Defense (TAMD) Capstone Requirements Document (CRD), develop text and mathematical definitions of ballistic missile SIAP attribute measures that are logical extensions to the initial SIAP attributes [DTIC Reference ADA 397215].
3. Define additional attributes related to timeliness and correctness applicable to ballistic missile analysis, based upon TAMD CRD requirements for ballistic missile attack EW, but not included in the initial SIAP attributes pertaining to air vehicles.

4. Identify any additional, lower-level metrics of possible relevance for assessment of the ballistic missile component of the SIAP as ballistic missile SIAP MOPs, and provide definitions.
5. Address any metrics implementation issues that are peculiar to the ballistic missile component of the SIAP.
6. Coordinate attributes with MDA and Services through the JTAMD process for comment.

FINDINGS

TAMD CRD EW attributes were well represented by the ballistic missile SIAP attributes. The full set of quantitative SIAP attribute measures and MOPs provided an adequate range of measurement possibilities to meet near-term SIAP assessment goals, covering the ballistic missile component for anticipated near-term test events. The level of mathematical rigor in the metric definitions provided a consistent extension of the approach taken for air vehicle SIAP metrics (SIAP SE TF Technical Reports 2001-001 through -003).

CONCLUSIONS

The metrics defined in this report provide a basis for quantitative assessments and comparisons of SIAP capabilities in connection with BMD. At the present stage of the SIAP metrics development effort, assessments and comparisons based on the ballistic missile metrics should prove useful in venues for which ballistic object reporting criteria are well defined. The groundwork laid in this report offers possibilities for extension into more complex domains.

RECOMMENDATIONS

Socialize Version 1.0 of this report within the joint community as a working document supporting near-term SIAP SE analytic and demonstration efforts involving a ballistic missile component. Extend and define the ballistic missile SIAP metrics in annual updates of this report, merging into the contents of other SIAP SE TF technical reports as applicable. Use the theoretical groundwork developed in this report and the experience gained through the use of these SIAP metrics to make recommendations for CRD development in connection with SIAP, BMD, and EW requirements.

TABLE OF CONTENTS

APPENDICES

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

This report continues the establishment, partially documented in Single Integrated Air Picture (SIAP) System Engineering Task Force (SE TF) Technical Reports 2001-001, -002, and -003, of a standard set of metrics for assessing the quality of the SIAP. The three earlier reports focused on the air vehicle component of the SIAP. Special assessment issues pertaining to the ballistic missile defense (BMD) component of the SIAP were pointed out briefly, but no attempt was made to cover these issues within the scope of the quantitative assessment framework represented by the SIAP attribute measures. Similarly, the list of SIAP Measures of Performance (MOPs) and Measures of Effectiveness (MOEs) introduced in SIAP SE TF Technical Report 2001-002 was also based on the implicit assumption that all objects of interest are unitary, and no measures specifically designed to support BMD evaluations or requirements were included. The present report addresses BMD issues exclusively, and provides an extension of the SIAP assessment methodology laid out in the earlier reports to cover these issues.

The problem of defining assessment objectives and a suitable set of metrics for BMD is intrinsically more complex than was the case for the air vehicle component of the SIAP. BMD is a relatively new warfare area, and the use of data links in support of BMD is still at an early and experimental stage. There is less past experience to draw upon than was the case for air vehicle SIAP assessments, and the direction of future development is less certain. Requirements for BMD are presently somewhat in flux (see Section 1.1 below). There are also technical complications arising from the fact that ballistic missiles may break up into multiple objects and spawn debris. Consequently, the assessment recommendations and metrics presented in this report must be regarded as somewhat provisional. The qualitative character of the ballistic missile SIAP attributes is expected to remain fairly stable with future developments. However, the definitions of the corresponding quantitative measures, the relative importance placed upon different metrics or different measurement approaches, and such distinctions as that drawn between attribute measures and MOPs, may all be subject to modification as more data is acquired and assessment goals become more precisely defined. All such modifications are to be incorporated in updates to this technical report, to appear at least annually starting with Version 2.0 later this year.

The remainder of this introduction describes the approach taken in adapting the SIAP attributes to the BMD context, and introduces a glossary of special terminology. Section 2 of this report provides a description of some key issues which complicate ballistic missile SIAP assessments (as compared with situations involving unitary air vehicles exclusively), and outlines the modifications of the SIAP assessment measures and procedures needed to resolve them. Section 3 details an amended list of SIAP attribute measures and MOPs appropriate for BMD-related evaluations. Discussion of outstanding issues

and plans for future work are included with the conclusions and recommendations in Section 4. Appendix A provides additional mathematical details on the new (and amended) metrics introduced in Section 3. Finally, Appendix B addresses the correspondences of the new metrics with the Theater Air and Missile Defense Capstone Requirements Document (CRD). A future update of this report will include an additional appendix covering correspondences with the BMD CRD (currently in draft).

1.1 CRD Guidelines

The philosophy guiding the development of SIAP metrics described in this report is essentially the same as that introduced for the air vehicle component in SIAP SE TF Technical Report 2001-001. In that earlier report, the basic attributes of *completeness*, *clarity*, *continuity*, *accuracy,* and *commonality* were defined, with one or more quantitative measures provided for each, so that the complete set was sufficient to represent all quantitative SIAP requirements specified in the TAMD CRD. (Variant definitions in the case of completeness, clarity, and accuracy were also introduced to cover Combat Identification CRD requirements.)

In extending the SIAP attributes to cover ballistic missile objects, a goal is to arrive at a set of attribute measures sufficient to represent all CRD requirements – in this case those pertaining to BMD and early warning (EW). As notions of "track picture quality" originally addressed in the context of air tracks are for the most part still relevant, the five SIAP attributes noted above are retained, augmented by two new ones. Additional measures for each attribute are defined in this report to cover the additional track information required for effective BMD and EW. The two SIAP attributes added to the original list are *timeliness* and *correctness*. As discussed in SIAP SE TF Technical Report 2001-001, "timeliness" is actually noted as a SIAP requirement in the TAMD CRD, but with quantitative aspects yet to be determined. *Correctness* is meant to capture the validity of ID-like information, such as booster type and object classification, which is particular to tracking of ballistic objects. The sense of *accuracy* is restricted in this report to accuracy of kinematic data. The precise sense in which *correctness* is used will be clarified by the correctness metrics defined in Section 3.1.

The ballistic missile SIAP metrics, as presented in this report, specifically answer only to Theater Ballistic Missile (TBM) EW requirements, although the general utility of the metrics is expected to be somewhat wider in scope. A Theater ballistic missile (TBM) EW Key Performance Parameter (KPP) and related requirements are documented in the TAMD CRD, including a number of "TBM EW attributes." These CRD-specified attributes are quantitative measures which may be grouped according to which of the SIAP attributes they represent, as shown in Figure 1.

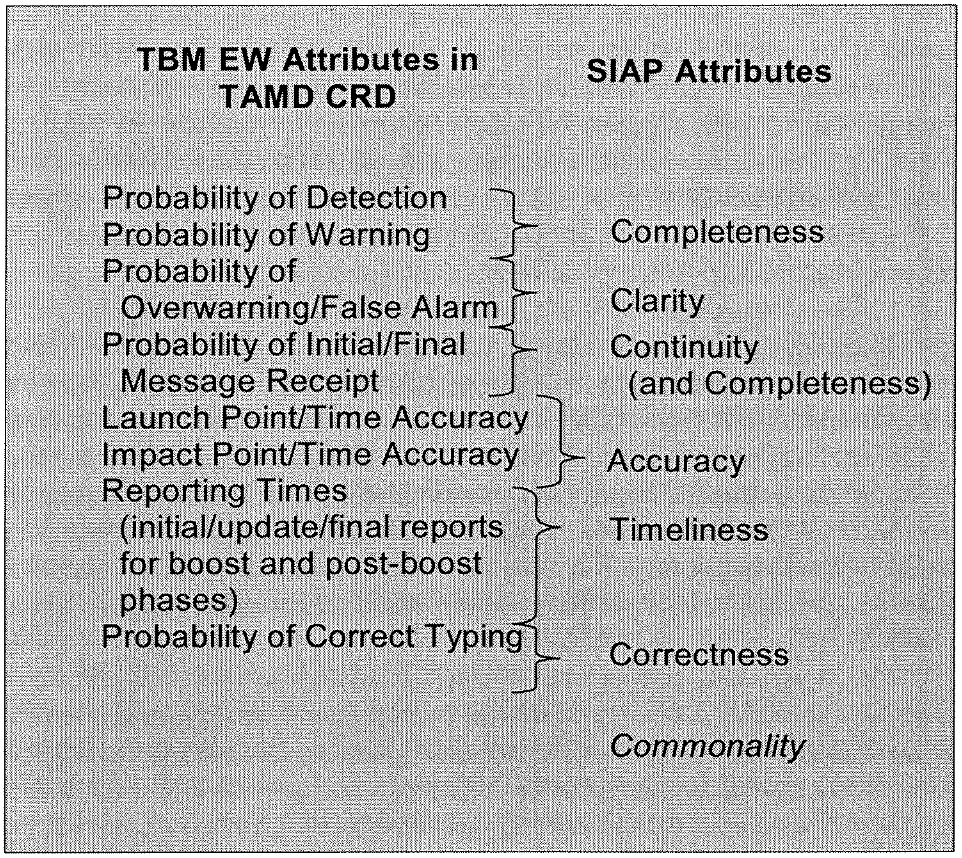

Figure 1. TAMD CRD and SIAP attribute correspondences.

As was the case in the air vehicle context, the SIAP attribute of *commonality* does not correspond to a specific CRD requirement, but may be considered part of the definition of a SIAP. The probabilities of initial and final message receipt are perhaps more directly measures of completeness than continuity, but if the initial and final messages are to be represented by (or relatable to) a single track identifier, a continuity requirement is implied.

The expanded set of SIAP attribute measures will be described in Section 3 of this report. Appendix B documents TAMD CRD KPPs and attributes in tabular form, and displays more detailed correspondences between TAMD CRD requirements and the SIAP attribute measures. It will be apparent from this tabular comparison that the EW KPP specified in the TAMD CRD corresponds to a set of particular instances of certain ballistic missile SIAP attribute measures.

Other BMD requirements are still being developed. As the draft BMD CRD matures, a similar set of relations to that depicted in Figure 1 will be developed for any BMD attributes defined therein. Detailed functional correspondences will be provided in Appendix C of a future update of this report.

It should be noted that the SIAP SE-led metrics development effort is progressing in concert with CRD development. The SIAP SE TF is interacting with the Joint Air and Missile Defense Organization (JTAMDO) and the Missile Defense Agency (MDA) to ensure that these parallel developments remain compatible while avoiding duplication of effort. The work described in this report is properly viewed as a contribution to the goal (common to all parties involved) of defining joint requirements covering SIAP, BMD, and EW, with due attention to the significant areas of overlap between these domains.

1.2 Glossary of Terms

The following specialized terminology is used throughout the remainder of this report. In cases where a term previously defined in SIAP SE TF Technical Report 2001-001 is re-defined in this report, the intent is to generalize the earlier definition so that the term applies to assessments involving either ballistic objects or air vehicles or both. Earlier definitions remain valid if "objects" are restricted as before to reportable air vehicle objects.

Area of Interest (AOI) – "that area of concern to the commander, including the area of influence, areas adjacent thereto, and extending into enemy territory to the objectives of current or planned operations. The area also includes areas occupied by enemy forces who could jeopardize accomplishment of the mission" (JP 1-02, 2001 and SIAP SE TF Technical Report 2001-001). The definition of an AOI will be further refined in each Common Reference Scenario (CRS), to bound the area within which SIAP attributes and other metrics will be evaluated. Typically the AOI for ballistic missiles can be expected to be different than that for air vehicles – including, for example, potential out-of-theater launch points, and regions of space through which ballistic threats may pass.

Reporting Criteria – discussed in Section 2.1 of this report. The intent is that the ballistic missile SIAP metrics should be compatible with any reporting criteria.

Object – any airborne or exo-atmospheric physical item of interest in the AOI. "Item of interest" depends on the operational context, and may be interpreted very broadly to include such items as ballistic missile debris, not normally construed as military threats.

Reportable Object – an object meeting reporting criteria. SIAP assessments based on the metrics defined in this report will only hold participants accountable for tracking reportable objects.

Object Thread – a parent-to-child string of objects (that is, each object, when it ceases to exist, spawns the next object in the

string), or a single object, satisfying further conditions elaborated in Section 2.1 of this report.

Currently Reportable Object Thread – an object thread that includes a reportable object existing at the current time.

Launch Point Estimate (LPE) or Determination – "estimate [of] the point on the earth's surface from which the missile was launched" (BMD Glossary, 1997). This report assumes a general situation in which an elliptical error region may be supplied. Also, launch time estimate, which relates to a particular LPE, is of interest in this report.

Impact Point Prediction (IPP) – "prediction of the point on the earth's surface where a specific reentry vehicle will impact.... The estimate includes the perturbing effects of the atmosphere and resulting uncertainties" (BMD Glossary, 1997). This report covers the general case of an elliptical error region. Also, predicted impact time, which relates to a particular IPP, is of interest in this report.

Assigned Track, Assigned LPE, Assigned IPP – in analysis of data from an exercise or simulation, a track, LPE, or IPP which meets a specified condition for association with an object, with a true launch point/time, or with a true impact point/time, respectively. Separate track-to-truth, LPE-to-true launch point/time, and IPP to true impact point/time assignments will be performed in accordance with Section 3.3 of this report, which adapts the general approach outlined in SIAP SE TF Technical Report 2001-003 to the ballistic missile setting. As with air vehicles, assignments are generally time dependent.

Large Caliber Rocket (LCR) – "an unguided, surface-launched, indirect-fire rocket with range equal to or greater than 20km that can be fired from a single or multiple launch platform" (TAMD CRD, 2001).

Valid Track – a track assigned to a reportable object.

Valid LPE – an LPE satisfying appropriate gating criteria (elaborated in Subsection 3.3.4.2 of this report) for positional and temporal proximity to the true launch point/time of a ballistic missile or LCR that has been launched.

Valid IPP – an IPP satisfying appropriate gating criteria (elaborated in Subsection 3.3.4.3 of this report) for positional and temporal

proximity to a true currently reportable object thread impact point/time.

Tracked Object – an object to which a track is assigned.

Covered Launch – a true launch point/time (of a ballistic missile or LCR that has been launched) to which an LPE is assigned.

Covered Impact – a true impact point/time of a currently reportable object thread, to which an IPP is assigned.

Track Switch – in the context of a gated unique optimal assignment of tracks to objects (v. SIAP SE TF Technical Report 2001-003 and Subsection 3.3.3 of this report), a change over time in the identity of the track assigned to a tracked reportable object (or object thread).

Track Break – in the context of a gated unique optimal assignment of tracks to objects, a change over time in the assignment status of a reportable object (or object thread) from tracked to untracked.

Declared Track, Declared LPE, Declared IPP – in analysis of data from an exercise or simulation, a track, LPE, or IPP which is subject to the assignment and scoring process for a particular participant on the basis that the participant assesses the track or IPP to be on a reportable object, or the LPE to be indicative of a true launch. Remote-only data must be declared if it is associated with a remote-only track classified by the source node as a reportable object type. Discussed further in Section 3.3.

Run – (1) for a live exercise, hardware in the loop, or operator in the loop event, simply a single experimental trial with data collected in near-real time over a fixed time period of interest; (2) the empirical data set so attained; (3) for a purely model-based simulation, a single execution of the simulation with fixed model start and end times and (if applicable) a fixed set of starting random number seeds; (4) the output data file(s) from a simulation run. In context, it is usually clear whether an actual event execution or merely a data set is meant, and for the purposes of this report the distinction is unimportant. In a pure simulation, typically many runs are executed, with all unconstrained inputs being randomized – a so-called "Monte Carlo Simulation." The basic definitions of all SIAP attribute measures are framed in terms of a single run. When many runs are executed Monte Carlo style, the ultimate measures of interest may be the averages over runs, or other statistics computed over runs.

2. SIAP ASSESSMENT ISSUES PARTICULAR TO BMD

2.1 Objects and Reporting Criteria

The ballistic missile SIAP attributes and MOPs depend on having a definition of a set of ballistic missile object types that are reportable. In general, an object is regarded as reportable if it is of interest for active defense, or if its impact point is of interest for passive defense, as determined by the user of this information. SIAP metrics are required to be fully compatible with scoring based on any such class of objects that a user might select (for active or passive defense purposes). Pending definitive guidance from users, the following classes will be considered as typical candidates for the class of reportable objects.

(1) A canonical illustrative list of objects able to cause substantial damage upon impact (or able to spawn such objects), consisting of: all active boosters, all post-boost objects containing a warhead, and all spent booster stages or segments containing a rocket motor.

(2) Any subset of the canonical list defined in (1), for example the subset of warhead-containing objects.

The user may also choose to specify several levels of reportability, for example, designating certain objects as reportable within a particular network, but only a subset of those objects as reportable outside of the network. The user taking this approach would also need to indicate which level of reporting would be used as a basis for scoring SIAP metrics. In this way, a user could collect information which might be useful in improving various functions within the network, without necessarily holding participants responsible in a formal scoring sense for reporting this information. The tracking of debris clouds containing no reportable objects (possibly of interest, but not for external reporting purposes) would be a typical application of this option. Most of the SIAP metrics defined in this report, and in particular the ballistic missile SIAP attribute measures defined in Section 3.1, are designed to assess the effectiveness of reporting on the basis of outside-of-network requirements such as EW. However, it should be possible to evolve the SIAP SE's assessment methodology to address intra-network functions as well, by exploiting the option for multiple levels of reporting just described, and defining function-specific MOPs based on the different levels. An example of a future intra-network requirement which might demand such an assessment approach would be the use of fused data from multiple sources in the generation of target-object maps (TOMs) for ballistic missile interceptors.

The user may also define circumstances under which a reportable object ceases to exist (or ceases to be reportable), for example, upon impact or upon a successful intercept. These circumstances could vary between different test venues, on the grounds that some venues might not support engagements or accurate scoring of intercepts.

It is expected that a more specific delineation of reportable objects for active and passive defense will eventually be worked out between the SIAP SE and other parties involved in the setting of BMD requirements (USJFCOM, JTAMDO, and MDA, at the minimum). Whenever such an understanding is established, it will be documented and its implications addressed in an update of this report.

Some of the definitions introduced in this report are dependent on the notion of a reportable *object thread*. A ballistic missile object thread is a string of parent-to-primary child ballistic missile objects (or a single object) that commences at launch or at the birth of a secondary child object and ends at the last time record of an object with no primary child (see Figure 1).

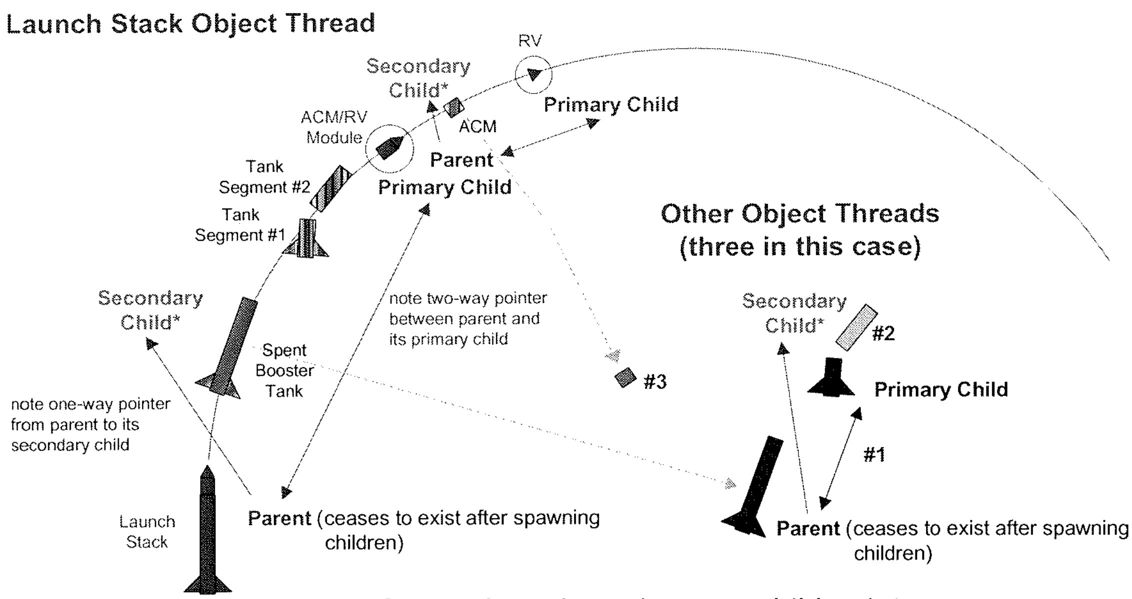

Figure 2. Object threads and parent-child pointers.

An object thread can be a single ballistic missile object or a series of objects. If it is a series of ballistic missile objects, the thread is inherited from the parent to its primary child object. The initial object in an object thread either has no parent (e.g., the launch stack), or is a secondary child object from an event that creates multiple child objects. Subsequent ballistic missile objects in an object thread are always primary children of their parent (e.g., the attitude control module/reentry vehicle (ACM/RV) stack is the primary child of the launch stack). Likewise, an ACM/RV stack spawns a primary child (the RV), a secondary child (the ACM), and separation debris (small debris objects are always secondary child objects). The RV is the final object in the launch stack object thread (note that the launch stack object thread is always traced through the series of ballistic missile objects that carry the warhead: launch stack, ACM/RV module, and finally the RV. The ACM initiates a new object thread and, in the example on the following chart, is the sole object in that thread. If the ballistic missile deploys

more than one warhead, then a recommended convention is to call the thread containing that last released warhead the launch stack thread, and the release of every other warhead starts a new reportable object thread.

There are reportable and non-reportable object threads. Reportable object threads can be launch stack object threads (always containing an object meeting reporting criteria). Launch stack object threads start with the launch stack and (for other than unitary ballistic missiles) propagate through the primary child object(s) carrying the warhead. Other reportable object threads are those that initiate as a secondary child that is a reportable object and may propagate through a subsequent primary child reportable object, if any. A non-reportable object thread is one that initiates as a secondary child that is a non-reportable object and may propagate through a subsequent primary child non-reportable object, if any. A non-reportable object never gives birth to a reportable object.

Since the approach taken to scoring LPEs in this report does not explicitly link LPEs to either tracks of objects or their impact points, a distinction between reportable and non-reportable launches is also presumed. A launch is reportable if it is a "true" launch, meaning a launch of a ballistic missile or LCR as defined in the glossary of Section 1.2. The essential criteria here, taken from the TAMD CRD, are that the launch be a surface launch, and that the rocket have a range of 20km or more and an indirect-fire capability.

Likewise, the approach adopted in this report to scoring IPPs does not explicitly link impact data to either the track of the impacting object or its launch data. In fact, the SIAP metrics for scoring IPPs are open to the possibility that an IPP might be derived from a conglomerate of track predictions, so that there would be no single track uniquely associated with the IPP report. Thus, a distinguished class of reportable future impacts is also presumed for scoring purposes. An impact point/time is regarded as reportable if the impacting object is part of a currently reportable object thread. This criterion implies, in particular, that the impending impact of a reportable object (i.e., with no further separation) is always reportable, while future impacts of objects spawned by a missile not yet launched are not reportable.

2.2 Measures Specific to Ballistic Missiles

In addition to tracks, ballistic missile SIAP attributes and MOPs are concerned with launch point/time estimates (LPEs), impact point/time predictions (IPPs), booster/missile type classification (BC), and post-boost object classification (PBC). Also, for ballistic missile tracks, track continuity over an entire parent-to-primary child object thread is of interest.

3. DEFINTIONS OF BALLISTIC MISSILE SIAP METRICS

This section introduces, in qualitative terms, the extended set of SIAP metrics to be used in BMD-related assessments. Most of the mathematical details are relegated to Appendix A.

Section 3.1 consists of verbal definitions for the ballistic missile SIAP attribute measures, grouped by attribute. These can be regarded as amendments to SIAP SE TF Technical Report 2001-001. Section 3.2 provides some possible ballistic missile SIAP MOPs, and similarly amends SIAP SE TF Technical Report 2001-002. The list of MOPs is not meant to be exhaustive. MOPs are intended for use in root-cause analysis and other analysis supporting SIAP critical experiments in certain test events (cf. SIAP Standard DMAP, 2002). The SIAP Analysis Team (SAT) will oversee refinement of the SIAP MOPs for suitability to these analysis efforts. Section 3.3 discusses some implementation issues (predominantly assignment issues) for ballistic missile SIAP metrics that did not arise in the simpler context of air vehicle metrics, and makes recommendations for extending the metrics implementation strategy of SIAP SE TF Technical Report 2001-003 to the ballistic missile context.

3.1 Ballistic Missile SIAP Attribute Measures

In the definitions which follow, *Method 1* and *Method 2* refer to two options for tracks-to-truth assignment, discussed below in Section 3.3.

Completeness

Object Track Completeness: Proportion of real objects meeting reporting criteria that are held as declared tracks at each scheduled scoring time in the scenario. This attribute measure can be rolled up over time for a particular participant, averaged across participants, or both, to yield a higher level measure or measures.

Launch Point Completeness: Proportion of real ballistic missiles that have been launched for which a valid launch point estimate is being provided. This attribute measure can be rolled up over time for a particular participant, averaged across participants, or both to yield a higher level measure or measures.

Impact Point Completeness: Proportion of objects meeting reporting criteria for which there is a valid impact point prediction. This attribute measure can be rolled up over time for a particular participant, averaged across participants, or both to yield a higher level measure or measures.

Clarity

Track Ambiguity: In a gated non-unique assignment using Method 1, the number of declared tracks that are assignable to real objects meeting reporting criteria, divided by the number of declared tracks actually assigned to reportable objects. This attribute measure can be rolled up over time for a particular participant, averaged across participants, or both to yield a higher level measure or measures. Track Ambiguity is currently undefined for truth-to-tracks assignment Method 2. If all the ballistic missile objects in the scenario are unitary ballistic missiles, then the Track Ambiguity attribute can be interpreted as redundant track mean ratio.

Spurious Track Mean Ratio: In a gated non-unique assignment using Method 2, the number of declared tracks that are unassignable to any real object, divided by the number of declared tracks. This attribute measure can be rolled up over time for a particular participant, averaged across participants, or both to yield a higher level measure or measures. Spurious Track Mean Ratio is currently undefined for truth-to-tracks assignment Method 1. As already noted in SIAP SE TF Technical Report 2001-001, test venues from which a reliable count of spurious tracks cannot reasonably be expected may not be required to score this measure.

LPE Ambiguity (Redundant LPE Mean Ratio): In a gated non-unique assignment, the number of declared launch point/time estimates that are assignable to real ballistic missile launches, divided by the number of valid declared LPEs. This attribute measure can be rolled up over time for a particular participant, averaged across participants, or both to yield a higher level measure or measures.

Spurious LPE Mean Ratio: In a gated non-unique assignment, the number of declared launch point/time estimates that are unassignable to real ballistic missile launches, divided by the number of declared distinct LPEs. This attribute measure can be rolled up over time for a particular participant, averaged across participants, or both to yield a higher level measure or measures.

IPP Ambiguity (Redundant IPP Mean Ratio): In a gated non-unique assignment, the number of declared impact point/time predictions that are assignable to true impact points/times of currently reportable object threads, divided by the number of valid declared IPPs. This attribute measure can be rolled up over time for a particular participant, averaged across participants, or both to yield a higher level measure or measures.

Spurious IPP Mean Ratio: In a gated non-unique assignment, the number of declared impact point/time predictions that are unassignable to true impact points/times of currently reportable object threads, divided by the number of declared distinct IPPs. This attribute measure can be rolled up over time for a

particular participant, averaged across participants, or both to yield a higher level measure or measures.

Continuity

Normalized Longest Duration Valid Track Segment on Distinct Reportable Objects: For a distinct real object meeting reporting criteria, the longest duration valid track segment without a track number switch or track break, normalized by the scenario lifetime of the object. This attribute measure can be aggregated across multiple reportable objects, aggregated across participants, or both to yield a higher level measure or measures.

Normalized Longest Duration Valid Track Segment on Reportable Object Threads: For a real primary-child thread of objects meeting reporting criteria, the longest duration valid track segment without a track number switch or track break, normalized by the scenario lifetime of the object thread. This attribute measure can be aggregated across multiple reportable objects, aggregated across participants, or both to yield a higher level measure or measures.

Accuracy

Track Accuracy: As assessed for a particular object meeting reporting criteria, the root mean squared error (RMSE) history in position, the RMSE history in velocity, the root sum squared average error (RSSAE) history in position, and the RSSAE history in velocity of the track assigned to that object compared to the truth states for that object. Position and velocity RMSEs can be rolled up over time for a particular participant, averaged across participants, or both to yield a higher level measure or measures for a particular object meeting reporting criteria.

LPE Position and Time Accuracy: As assessed for a particular real ballistic missile launch, the RMSE in geographic position (measured in terms of geodesic path length on the WGS-84 ellipsoid) and the RMSE in time of the declared LPE assigned to that launch point/time. These RMSEs can be rolled up over time for a particular participant, averaged across multiple ballistic missile launches, averaged across participants, or aggregated in combinations of these to yield a higher level measure or measures.

IPP Position and Time Accuracy: As assessed for a reportable object thread true impact, the RMSE in geographic position (in terms of geodesic path length on the WGS-84 ellipsoid) and the RMSE in time of the declared IPP assigned to that impact point/time. These RMSEs can be rolled up over time for a particular participant, averaged across reportable thread impact points, averaged across participants, or aggregated in combinations of these to yield a higher level measure or measures.

Timeliness

Track Initiation Delay: Delay from birth of a distinct reportable ballistic missile object (i.e., launch time for a unitary ballistic missile or a full launch stack, otherwise time of separation, release, breakup, segmentation, etc., that gives rise to the distinct reportable object), until the first time at which the object has a valid declared track. This attribute measure can be aggregated across reportable objects of particular types/classifications of interest, across participants, or both to yield a higher level measure or measures.

Launch Point Estimate (LPE) Delay: Delay from launch of a ballistic missile or LCR until the first launch point estimate is assigned to that true launch. This attribute measure can be aggregated across multiple ballistic missile launches, aggregated across participants, or both to yield a higher level measure or measures.

Booster Burnout Estimate Delay: Delay from burnout of the final booster stage of a ballistic missile until the first valid booster burnout estimate with state vector is provided. This attribute measure can be aggregated across multiple ballistic missiles, aggregated across participants, or both to yield a higher level measure or measures.

Impact Point Prediction (IPP) Warning Time: For a particular reportable object thread impact, the true time of impact minus the first time a ground impact point prediction is assigned to that true impact. This attribute measure can be aggregated across multiple reportable objects from one or more ballistic missiles, aggregated across participants, or both to yield a higher level measure or measures.

Correctness

Booster Typing Correctness: As assessed for a particular real ballistic missile launch, the match between the booster type probability vector (or single type declaration) reported with a valid LPE and the true booster type for that launch. This attribute measure can be rolled up over time for a particular participant, averaged across multiple ballistic missile launches, averaged across participants, or aggregated in combinations of these to yield a higher level measure or measures.

Post-Boost Object Classification Correctness: As assessed for a particular real post-boost object meeting reporting criteria, the match between the reported object classification probability vector (or single classification declaration) associated with a valid track and the true classification of the post-boost object. This attribute can be rolled up over time for a particular participant, averaged

across participants, or both to yield a higher level measure or measures for a particular object meeting reporting criteria.

IPP Object Classification Correctness: As assessed for a particular reportable object thread impact, the match between the impacting object classification probability vector (or single classification declaration) associated with the declared IPP assigned to that impact point/time and the true classification of the impacting object. This attribute measure can be rolled up over time for a particular participant, averaged across reportable thread impact points, averaged across participants, or aggregated in combinations of these to yield a higher level measure or measures.

Commonality

Track Commonality: Ratio of the number of valid declared tracks held by all participants simultaneously, with consistent track numbers, kinematic data, and typing/classification data, to the total number of valid declared tracks.

LPE Commonality: Ratio of the number of true launch points/times, with consistent booster typing/classification data, for which all participants hold a declared LPE assigned to the true launch, to the total number of valid declared LPEs.

IPP Commonality: Ratio of the number of true impact points of reportable objects, with consistent impacting object classification data, for which all participants hold a declared IPP, to the total number of valid declared IPPs.

3.2 Ballistic Missile SIAP MOPs

Track Continuity

Cumulative Switches of Tracks on Distinct Reportable Objects: For distinct real objects meeting reporting criteria, the cumulative number of switches of tracks (not counting breaks) by time t into the scenario for particular objects, and summed across all reportable objects. This MOP can also be averaged across participants to yield a higher level measure.

Cumulative Switches of Tracks on Reportable Object Threads: For real primary-child threads of objects meeting reporting criteria, the cumulative number of switches of tracks (not counting breaks) by time t into the scenario for particular threads and summed across all reportable object threads. This MOP can also be averaged across participants to yield a higher level measure or measures.

Cumulative Broken Tracks on Distinct Reportable Objects: For distinct real objects meeting reporting criteria, the cumulative number of breaks of tracks by time t into the scenario for particular objects, and summed across all reportable objects. This MOP can also be averaged across participants to yield a higher level measure.

Cumulative Broken Tracks on Reportable Object Threads: For real primary-child threads of objects meeting reporting criteria, the cumulative number of breaks of tracks by time t into the scenario for particular threads, and summed across all reportable object threads. This MOP can also be averaged across participants to yield a higher level measure.

Accuracy

Track Covariance Consistency: As assessed for a particular object meeting reporting criteria, the mean normalized six-state Chi-squared statistic of the track assigned to that object. This MOP can be rolled up over time for a particular participant, averaged across participants, or both to yield a higher level measure or measures for a particular object meeting reporting criteria.

LPE Position and Time Uncertainty Consistency: TBR (As assessed for a particular real ballistic missile launch, the mean Chi-squared statistic of the LPE time estimate and the proportion of occurrences in which the true position error of the declared LPE is within the 95th percentile uncertainty areas reported with the LPE. This MOP can be aggregated across multiple ballistic missile launches, averaged across participants, or both to yield a higher level measure or measures.)

IPP Position and Time Uncertainty Consistency (Deterministic Future Truth Approach): TBR (As assessed for a particular real object meeting reporting

criteria, the mean Chi-squared statistic of the IPP time estimate and the proportions of occurrences in which the true position error of the declared IPP is within the 95th percentile uncertainty area reported with the IPP. This MOP can be aggregated across multiple reportable object threads, averaged across participants, or both to yield a higher level measure or measures.)

Pair-wise Cross-Platform Commonality History

Ratio of Non-Common Network Track Numbers: Number of active network track numbers that are different (additions or deletions) between pairs of participating units, divided by the number of network track numbers in the union of the two network track files. This MOP can be aggregated across all participant pairs to yield a higher level measure.

Track State Estimate Differences: The Euclidean differences between position and velocity state estimates of tracks held by pairs of participating units, for tracks with the same active network track number. This MOP can be rolled up over time for a particular pair of participants, averaged across all participant pairs, or both to yield a higher level measure or measures.

Ratio of Non-Common Launch Points: Number of covered true ballistic missile launches that are different (additions or deletions) between declared LPEs of pairs of participating units, divided by the number of true ballistic missile launch covered in the union of the LPEs. This MOP can be aggregated across all participant pairs to yield a higher level measure.

LPE Differences: For pairs of participating units, the geographical position and time differences between LPEs held by the pair that are assigned to the same true ballistic missile launch. This MOP can be rolled up over time for a particular pair of participants, averaged across all participant pairs, or both to yield a higher level measure or measures.

Ratio of Non-Common Impact Points: Number of covered true reportable object thread impacts that are different (additions or deletions) between declared IPPs of pairs of participating units, divided by the number of true reportable object thread impacts covered in the union of the IPPs. This MOP can be aggregated across all participant pairs to yield a higher level measure.

IPP Differences: For pairs of participating units, the geographical position and time differences between IPPs held by the pair that are assigned to the same true reportable object thread impact. This MOP can be rolled up over time for a particular pair of participants, averaged across all participant pairs, or both to yield a higher level measure or measures.

Booster Type Estimate Differences: TBD (For pairs of participating units, the norm of the difference vector between boost type probability vectors, divided by the sum of the norms of the boost type probability vectors.) This MOP can be rolled up over time for a particular pair of participants, averaged across all participant pairs, or both to yield a higher level measure or measures.

Post-Boost Object Classification Differences: TBD (For pairs of participating units, the norm of the difference vector between post-boost object classification probability vectors, divided by the sum of the norms of the post-boost object classification probability vectors.) This MOP can be rolled up over time for a particular pair of participants, averaged across all participant pairs, or both to yield a higher level measure or measures.

System Loading

Communications Data Loading: Sum of data rates input from all platforms into the communications function for distribution to other remote platforms.

3.3 Ballistic Missile Metrics Implementation Issues

This subsection addresses a number of implementation issues, primarily assignment to truth and related issues, which are inherent to the truth-centric nature of the ballistic missile SIAP metrics. As already noted in the Introduction, the approach taken in this report requires separate assignment determinations for tracks, LPEs, and IPPs.

The recommendations of this section, particularly with regard to assignment algorithms, should be regarded as provisional. They should be suitable for enabling near-term analytical work dependent on evaluation of SIAP metrics. However, the recommended assignment procedures have yet to be thoroughly tested, and may require revision on the basis of some initial metrics analysis following planned events/simulations over the coming year.

3.3.1 Outline of Truth-Centric Assessment Approach

Like the air vehicle SIAP attributes and MOPs, the ballistic missile SIAP attributes and most ballistic missile SIAP MOPs are truth centric rather than track centric. The truth centric approach for ballistic missile SIAP attributes and MOPs is summarized as follows:

- Independently know true object states and object classification over time, as well as the true ballistic missile launch points/times and reportable object thread impact points/times. This is straightforward in simulations, but requires instrumentation and/or designated truth data collection sensors in live tests and exercises.
- Schedule performance scoring times throughout interval of interest. Scheduled times can be different for scoring tracks, LPE/BC, IPP, and PBC performance. In Monte Carlo simulation, scheduled scoring times for the same performance metrics must be aligned across multiple runs to support statistical analysis. Also, for the purpose of isolating key event times, ballistic missile SIAP-related data should be scored whenever a new firm track, a new LPE, or a new IPP is generated or otherwise obtained.
- At scheduled and first-time scoring times, the ballistic missile SIAP-related estimates/predictions and error covariance/uncertainty representation data are collected to be scored. For ballistic missile tracks, the data of interest are the position/velocity states and their error covariance (if available) at the last track update time or predicted to the scoring time, depending on the selected time alignment scheme. For LPEs, the data of interest are the launch position/time estimates and uncertainty representation for each perceived parent ballistic missile. For IPPs, the data of interest are the ground impact point/time predictions and uncertainty representation for each perceived unique object satisfying reporting criteria. For booster/missile typing classification and for post-boost object classification, the preferred data to collect (if available) are vectors of assessed classification probabilities. If probabilistic classification

data are unavailable, then the reported booster type or object class is the datum of interest.

- The state estimates/predictions are then paired to truth with a suitable 2-data-set assignment procedure and the ballistic missile SIAP-related data are then scored according to the performance attributes and MOPs. Assignment of track states to truth object position/velocity states is used to score ballistic missile tracks and post-boost object classifications assessments. Assignment of LPEs to true launch points/times is used to score LPEs and booster/missile type classifications. Assignment of IPPs to true currently reportable object thread impact points/times is used to score IPPs and impacting object classifications.

3.3.2 Scoring Precedence

Scoring precedence rules for air tracks (in SIAP SE TF Technical Report 2001-003) were based on the availability of a "central track store" (CTS) for each participant, to act as a primitive track base from which certain groups of tracks would be selected for independent assignment and scoring. It is desirable to preserve several aspects of this down-selection (for example, assigning all local tracks in an independent pass from the remote tracks) for ballistic missile SIAP assessments. However, the CTS (or its analog) is not as well-defined when all possible sources of ballistic object data are taken into consideration.

The following guideline has been adopted for selecting the data to be scored in assessing the ballistic missile component of the SIAP. Any store of data, including LPE and IPP data, which is displayable *on the same display* as would normally be used for a TADIL, is to be regarded as scorable, excluding remote mutual track data. This simple guideline may not be sufficient to determine the appropriate data extraction points for every system, but system-specific data extraction issues will have to be addressed individually as they arise. The guideline is provided so that the prescription of data extraction points will be based on consistent criteria.

The scoring precedence guideline is to be regarded as a default. Special-purpose SIAP assessments may be undertaken for other classes of data, along lines suggested in SIAP SE TF Technical Report 2001-003.

3.3.3 Tracks-to-Truth Assignment

3.3.3.1 Comparison of Assignment Methods

In ballistic missile debris environments, assigning tracks only to truth objects meeting reporting criteria may not provide sufficient pattern matching for good metrics assignment performance. However, in some networks and operator display environments, there may be an intention only to provide tracks on reportable objects. If just a few tracks are assigned against a large set of truth

objects including those that are not reportable objects, there is increased chance that tracks will be assigned to non-reportable objects and consequently scored poorly with respect to track, IPP, and PBC attributes and MOPs. For these and the other reasons noted below, two parallel methods are being implemented for experimentation with ballistic missile tracks-to-truth assignment algorithms, at least in simulation environments. As explained below, Method 1 will normally be regarded as the default for scoring the SIAP attributes.

Method 1: Make it the responsibility of the holder of the ballistic missile tracks to down-select the tracks to those it assesses to be on reportable objects (the set of "declared" tracks as defined in Section 1.2), then assign the smaller set of retained tracks just to active truth objects meeting reporting criteria. The advantages of Method 1 include its similarity to the method used to assign air vehicle tracks to truth, reduced computational burden from the metrics assignment perspective, and, when the ballistic missiles remain unitary objects, retention of the familiar definition of redundant/dual tracks. The disadvantages include there being insufficient pattern in some scenarios to provide optimal tracks-to-truth assignment, especially with attendant sensor biases; sources or holders of ballistic missile tracks may inadvertently submit the wrong tracks or try to cover all possibilities by submitting many extra tracks, and the method does not reveal tracks that are spurious in an absolute sense.

Method 2: Assign all tracks to all active truth object representations, but only score tracks assigned to objects meeting reporting criteria. The advantages of Method 2 include good metrics assignment pattern matching and identification of tracks that are spurious in an absolute sense. The disadvantages include a potentially large computation task to perform the tracks-to-truth assignment and the conceptual issue of how to assess redundant/dual tracks.

Because Method 2 may not be feasible in certain venues (particularly live evaluations), Method 1 will be the default for scoring all SIAP attribute measures except for the Spurious Track Mean Ratio. For venues in which it is feasible to implement both assignment methods, Method 2 will always be used for Spurious Track Mean Ratio, and Method 1 will be the default for other metrics.

For both methods, SIAP-related estimates/predictions held by all platforms will be scored. Tracks, LPEs/booster type estimates, IPPs, and post-boost object classifications must be created at or distributed to each participating platform. Performance metrics will be computed for each participating platform and for pairs of participating platforms. Communications loading can be scored as well.

3.3.3.2 Time Aligning

Time aligning for ballistic missile tracks-to-truth assignment will be done in accordance with the default procedure for air tracks described in SIAP SE TF Technical Report 2001-003. Briefly, truth data is to be interpolated to the last

track update time (for local track data) or to the time stamp on the track (for remote track data), and the interpolated truth compared with the track data for assignment cost and gating purposes. Since all J3.6 messages carry time stamps, one of the reservations about using this method with J3.2 messages – that there might be unaccounted-for latencies in the track data (SIAP SE TF Technical Report 2001-003) – does not apply in the present context.

3.3.3.3 Assignment Algorithm

An automated tracks-to-truth assignment algorithm, similar but not identical to the one introduced for air tracks in SIAP SE TF 2001-003, is being developed for use in ballistic missile SIAP assessments.

The following procedural steps follow the basic outline of the algorithm used for air tracks:

Separately for node m at each scheduled scoring time,

Step 1a: Down select to the set of current local tracks held by node m that have precedence for scoring (based on the guideline stated in Section 3.3.2)

Step 1b: When using Assignment Method 1, down select further to the subset of local tracks with scoring precedence that node m declares to be on reportable objects, by whatever the local rule is.

Step 2: Discard from consideration at this scoring time all remote tracks held by node m which that node's track correlation function currently assesses to be mutual with node m local tracks.

Step 3a: Down select the tracks retained in Step 2 to the set of current non-mutual remote tracks held by node m that have precedence for scoring.

Step3b: When using Assignment Method 1, down select further to the subset of non-mutual remote tracks with scoring preference (1) that the remote source classifies as reportable objects and (2) for which the remote source provides no object classification. If the network from which the remote track is received employs a probability vector method for conveying object classification information, then for Assignment Method 1 down select to the subset that the recipient node considers reportable, by whatever the local rule is.

Step 4: Collect the tracks retained in Steps 1-3 into a single base of tracks held by node m which are eligible for assignment. Assign eligible tracks to truth objects (reportable truth objects if using Method 1) using a gated unique optimal assignment algorithm (GUOA). The results of this assignment step are used to score all system metrics for node m whose calculations require unique assignment (see Appendix A for computational details).

Step 5: Separately from the unique assignments determined in Step 4, identify the set of eligible tracks in Step 4 which satisfy gating criteria for association with truth objects (reportable truth objects if using Method 1). This may result in nonunique associations of tracks with objects. The results of this step are used to score the remaining metrics (those not dependent upon unique assignment – see Appendix A for details).

Repeat the above steps separately for each node being evaluated.

The following features of the assignment algorithm under development will differ from those of the air track assignment algorithm.

- The air track assignment algorithm applies a GUOA step to local tracks with precedence, followed by a separate GUOA step for each group of non-mutual remote tracks with precedence originating from a single remote source. A rationale for this use of separate assignments for distinct source nodes is given in SIAP SE TF Technical Report 2001-003. It is felt, however, that the presence of numerous debris tracks (i.e., tracks on non-reportable objects) in the ballistic missile context may invalidate the rationale invoked for air track assignments. Thus, although both procedures are based on a GUOA step, there is a significant difference in how the GUOA is applied. Until some preliminary results are analyzed (and even for the more familiar setting of air track assignment, analysis of assignment algorithms is at an early stage), it remains unclear which, if either, procedure yields more reliable results. Likewise, it is unclear at this point whether there is any particular drawback to approaching the assignment problem differently in the two different settings. This will be a major subject for investigation as assignment algorithms are implemented.

- For ballistic missile assessments, the cost function for the GUOA must be based on full six-dimensional track and truth data, the altitude components being equal in importance to the horizontal components.

- It will be convenient to compute assignment cost in an earth-centered, earth fixed (ECEF) Cartesian frame, as the data is encoded in the J3.6 message, rather than a track-centered local East-North-Up (ENU) frame as was recommended for air vehicle tracks.

- Unlike air vehicle tracks which tend to reach a steady-state accuracy very quickly, ballistic missile track accuracy typically changes several orders of magnitude over time, so it will not be suitable to assume a representative fixed accuracy per sensor type in the tracks-to-truth assignment cost and gating. If ballistic missile track error covariance could be relied upon to be consistent and always available (including in remote-only tracks) then a Mahalanobis distance over the 3-D position and velocity state estimates

would be a preferred cost and gate function. Unfortunately, initial experiments suggest that track error covariance is not a consistent measure of track error when data association is challenging. Also, the error covariance words for the J3.6 message are not transmitted with every track update, so there is no assurance that these data are available for remote-only tracks. For these reasons, final recommendations for the tracks-to-truth assignment cost and gating are not yet available. A prototype algorithm is currently using the Euclidean position difference squared, but this is less than ideal.

- Implementation of a hysteresis factor is still being investigated. As with air vehicle tracks-to-truth assignment, applying hysteresis may help to reduce unwarranted switching.

3.3.4 Assignment and Gating Criteria for LPEs and IPPs

The procedures followed for LPEs and for IPPs are slightly different (although similarly motivated) from the track assignment procedure described in Subsection 3.3.3.

3.3.4.1 General LPE and IPP Assignment Procedures

Separately for node m at each scheduled scoring time,

Step 1a: Down select to the set of current LPEs and IPPs generated with local track data held by node m that have precedence for scoring.

Step 1b: When using tracks-to-truth Assignment Method 1, down select further the set of IPPs generated with local track data with scoring precedence to the set of IPPs that node m declares to be on reportable object threads.

Step 2: If node m has an LPE or IPP correlation function, discard from consideration at this scoring time all remote LPEs and IPPs held by node m which that correlation function currently assesses to be mutual with LPEs and IPPs generated with node m local track data.

Step 3a: Down select to the set of current (non-mutual) remote LPEs and IPPs held by node m that have precedence for scoring.

Step3b: When using tracks-to-truth Assignment Method 1, down select further the set of (non-mutual) remote IPPs with scoring preference to the subset (1) that the remote source classifies as being IPPs on reportable object threads and (2) for which the remote source provides no object classification. If the network from which the remote IPP is received employs a probability vector method for conveying object classification information, then for tracks-to-truth Assignment

Method 1 employ the same criteria used in Step 1b to down select further the remote-only IPPs.

Step 4: Separately assign all retained LPEs and IPPs generated with local track data to true launches and reportable object thread impacts, respectively, using a first pass through a gated unique optimal assignment (GUOA) algorithm, followed by a second pass on not yet assigned LPEs and IPPs generated with local track data, through a gated independent nearest neighbor assignment (INNA).

Step 5a: Separately assign retained non-mutual remote LPEs and IPPs from remote source node n using the same procedure and formulas as for retained LPEs and IPPs generated with local track data.

Step 5b: If there are more sources of retained non-mutual remote LPEs or IPPs, increment n and return to Step 5a.

Step 6: Merge the separate assignments into complete lists of LPEs-to-truth and IPP-to-truth assignments for LPE and IPP data held at node m.

Repeat the above steps separately for each node being evaluated.

3.3.4.2 LPE Gating and Assignment Procedure

LPEs are gated and assigned directly to true ballistic missile launch events. The LPE gating and assignment costs are to be computed as follows:

At each LPE scoring time of each run, obtain from the participating unit the set of declared LPEs, each consisting of a launch time estimate t_i with variances $\sigma_{t\_i}^2$ and launch point estimate latitude and longitude coordinates (ϕ_i, λ_i) with 95th percentile ellipse major and minor axes, a_{maj_i} and a_{min_i}, and azimuth orientation of the major axis β_I, measured from local East. (If no value for $\sigma_{t\_i}^2$ is provided/available, assume $\sigma_{t\_I} = G_t / 3$, where G_t is defined below.)

First, gate LPE i with respect to true launch time j by $|t_i - t_j| < G_t$ (e.g., G_t = 30.0 sec).

If the gating above is passed, compute the WGS-84 ellipsoid minimum surface path (geodesic) distance $R_{i,j}$ and azimuth angle $\alpha_{i,j}$, measured from local east, from LPE i to true launch point j (using the Vincenty method, for example), then gate the LPE/true launch combination with respect to maximum surface distance by $R_{i,j} < G_r$ (e.g., G_r = 20 km).

Assuming the distribution of launch time estimate error and the underlying LPE position estimate error is close to trivariate normal, for LPE/true launch combinations that are not precluded by time estimate or surface distance gating,

compute the one sigma axis dimensions, σ_{maj_i} and σ_{min_i}, by multiplying a_{maj_i} and a_{min_i} by 0.4085, then gate the LPE time/position by

$$C(i,j) = \left[\frac{t_i - t_j}{\sigma_{t\_i}}\right]^2 + \left[\frac{R_{i,j}\cos(\alpha_{i,j} - \beta_i)}{\sigma_{maj\_i}}\right]^2 + \left[\frac{R_{i,j}\sin(\alpha_{i,j} - \beta_i)}{\sigma_{min\_i}}\right]^2 < G \qquad (1)$$

(e.g., G = 30.67, corresponding to a 0.99999 probability gate for a chi-squared statistic with three degrees of freedom).

Across all the LPE/true launch combinations (i,j) that are not gated out by any of the above criteria, perform an optimal unique assignment with the C(i,j) values from the above formula, setting the guard value equal to G (cf. SIAP SE TF Technical Report 2001-003, Section 3.3).

3.3.4.3 IPP Gating Procedure

IPPs are gated and assigned directly to true reportable object thread impacts. The IPP gating and assignment costs are to be computed as follows:

At each IPP scoring time of each run, obtain from the participating unit the set of declared IPPs, each consisting of an impact time prediction t_i and impact point prediction latitude and longitude coordinates (ϕ_i, λ_i) with 95th percentile ellipse major and minor axes, a_{maj_i} and a_{min_i}, and azimuth orientation of the major axis β_I, measured from local East.

First, gate the IPP i with respect to true primary-child threat impact time j by $|t_i - t_j| < G_t$ (e.g., $G_t = 60$ sec).

If the gating above is passed, compute the WGS-84 ellipsoid minimum surface path (geodesic) distance $R_{i,j}$ from the IPP i to the true primary-child thread impact location j, then gate the IPP/true primary-child thread impact combinations with respect to maximum surface distance by $R_{i,j} < G_r$ (e.g., Gr = 40 km).

Finally, gate the IPP time/position by

$$C(i,j) = \left[\frac{t_i - t_j}{G_t}\right]^2 + \left[\frac{R_{i,j}}{G_r}\right]^2 < G \qquad (2)$$

(e.g., G = 2.0).

Across all the IPP/true reportable object thread impact combinations (i,j) that are not gated out by any of the above criteria, perform an optimal unique assignment with the C(i,j) values from the above formula, setting the guard value equal to G (cf. SIAP SE TF Technical Report 2001-003, Section 3.3).

3.3.5 Outstanding Assignment Issues

- For the initial ballistic missile SIAP attributes implementation, there will be no special treatment of unresolved closely spaced objects (UCSO), nor cluster tracking constructs. Precise definition of USCO and clusters are open issues that lack BMD community consensus on mathematical formulations. Consequently, declared tracks will be assigned to the entire set of truth object representations involved in the invoked metrics assignment method.
- If a track source outputs aggregate tracks, aggregate LPEs, or aggregate IPPs with identified strength N, the scoring procedure is not yet formalized. In part, this is because guidance from USJFCOM on aggregate tracking in the BMD context is still forthcoming. (CINCUSJFCOM J8 Memorandum of 29 June, 2001). In the interim until a procedure is formalized, the intent is to ensure that SIAP metrics are adaptable to whatever future requirements for aggregate tracking might be defined. If it is necessary to score aggregate tracks before formal guidance is available, the metrics interface will replicate those tracks, LPEs, or IPPs N times and proceed with assignment to truth and scoring as for non-aggregated outputs.

Because of the need for further experimentation with assignment algorithms, final recommendations on issues related to assignment may be deferred until sufficient data is collected and analyzed.

4. CONCLUSIONS AND RECOMMENDATIONS

It is expected that the assessment approach and the specific SIAP metrics defined in this report will meet the near-term ballistic missile SIAP assessment needs of the SIAP SE – primarily defined as limited HWIL and simulation activity through the remainder of calendar year 2002. The envisioned HWIL activity, under development as of this writing, involves experimentation using straightforward pre-defined reporting criteria for ballistic objects. The envisioned simulation activity will involve mostly Monte Carlo simulation environments, within which prototypes of the assessment approach described in this report have already been explored with some success. It is not expected that the gaps in the description of reporting criteria or tracks-to-truth assignment will present any major barriers to these near-term assessment efforts.

This report will be updated on at least an annual basis, with an update to Version 2.0 scheduled for early in 2003. Firmer recommendations as to tracks-to-truth assignment algorithm should be available by that time, and clearer understandings of reporting criteria and formation/aggregate tracking requirements will be incorporated as they become available. The goal for Version 2.0, apart from providing a firmer stance on some of the points of the present document, will be to answer to the more demanding needs for ballistic missile SIAP assessments through calendar year 2003.

5. REFERENCES

Ballistic Missile Defense (BMD) Glossary. Version 3.0. (1997, June). Washington, DC: Ballistic Missile Defense Organization.

CINCUSJFCOM J8 Memorandum (2001, June 21). Subj: Single Integrated Air Picture System Engineer Block 0 Support Request. Norfolk, VA: USJFCOM J85.

Joint Publication (JP) 1-02: Department of Defense Dictionary of Military and Associated Terms. (2001, April 12). Washington, DC: Joint Chiefs of Staff.

SIAP SE TF Standard DMAP. (2002, March draft)

SIAP SE TF Technical Report 2001-001: Single Integrated Air Picture (SIAP) Attributes. (2001, June). DTIC Reference ADA 397215. Arlington, VA: SIAP SE TF.

SIAP SE TF Technical Report 2001-002: Single Integrated Air Picture (SIAP) Measures of Effectiveness (MOEs) and Measures of Performance (MOPs). (2001, October). DTIC Reference ADA 397221. Arlington, VA: SIAP SE TF.

SIAP SE TF Technical Report 2001-003: Single Integrated Air Picture (SIAP) Metrics Implementation. (2001, October). DTIC Reference ADA 397225. Arlington, VA: SIAP SE TF.

Theater Air and Missile Defense Capstone Requirements Document (TAMD CRD). (2001, March). Norfolk, VA: U.S. Joint Forces Command.

APPENDIX A

Mathematical Details Pertaining to Definitions of Ballistic Missile SIAP Metrics

Introduction

This appendix documents the mathematical formulae by which the ballistic missile SIAP metrics are to be calculated.

Each of the ballistic missile SIAP metrics may be averaged in various ways to obtain "roll-up" metrics. Explicit formulae for roll-up averages are provided only for the ballistic missile SIAP attribute measures. Since averaged measures (or, in certain cases, percentile values) of the SIAP attributes will be used to assess compliance with CRDs, it is essential that these roll-up procedures be agreed upon in advance. The recommended procedure is not the same for every SIAP attribute measure. For those measures which the TAMD CRD invokes for requirements pertaining to particular events or applicable at particular instants of time (see Appendix B, Tables 2-3), the roll-up should be the measure evaluated for that particular event or time, averaged over participants (and objects or events as appropriate). In other cases where a TAMD CRD requirement may be interpreted as applying to an attribute measure in a time-averaged sense, the recommended roll-up metric is an appropriately weighted average over time and participants (and objects or events as appropriate). In cases where the TAMD CRD specifies a particular percentile value of an attribute measure, a joint sample cumulative distribution of all measurements (over all participants, times of interest, etc.) is recommended. Any desired percentile value can easily be extracted from the distribution. In cases where the current TAMD CRD (2001) is unclear as to the level of averaging, an interpretation has been made on the basis of current SIAP SE TF / JTAMDO draft recommendations to USJFCOM for the next revision of the TAMD CRD.

To accommodate evaluation of the metrics in simulation environments where multiple runs may be executed (Monte Carlo simulations), averages over runs (or cumulative sampling over runs) are also embedded in the roll-up procedures.

It is expected that the user will define the time intervals over which each metric is to be evaluated (and averaged, if appropriate), as well as a schedule of scoring times. The intervals of evaluation/averaging need not be the same for every metric. In some cases, a TAMD CRD requirement may imply a time period of interest, which may differ among participants or among different objects being assessed (for example, IPP accuracy requirements specify a certain time period before impact for each reportable object). The mathematical notation used in this Appendix assumes a pre-defined sequence of scoring subintervals (the k^{th} such subinterval being of duration Δt_k), and a pre-defined scheme for selecting a

scoring time t_k from the kth subinterval. The sequence of subintervals and the selection of scoring times may be different for the scoring of tracks, LPEs, and IPPs, but should otherwise be consistently defined for all applications of the ballistic missile SIAP metrics in the context of a particular CRS. Event-driven scoring may also be specified independently for certain critical events (such as track initiation or first LPE generation), to cover requirements applying to those specific events.

An allowance for the possibility that the time intervals of interest may be metric-specific (or object-specific, etc.) is embedded into the formulation of the relevant averages, as demonstrated by the following. Suppose a single weighted average value is required for a metric $V_{i,m,n}(t_k)$, the metric being defined in a way that is specific to the object (i), participant (m), scoring time (t_k) and run (n) being assessed. For some values of the arguments, the metric may possibly be undefined (perhaps, for example, if i designates an object on which participant m holds no assigned track). Let the values designated by $W_{i,m,n,k}$ be a set of nonnegative weights which are meaningful and appropriate for all values of (i,m,n,k) for which $V_{i,m,n}(t_k)$ is defined. In the formulae for specific attribute measures which follow in the remainder of this Appendix, the weighting may be by number of reportable objects or events, number of participating units, or reportable object lifetimes, as applicable. By convention, set $W_{i,m,n,k}$ equal to 0 if *either* (1) $V_{i,m,n}(t_k)$ is not defined, or (2) the scoring time t_k is not contained in the time interval(s) of interest for this metric as considered for object i, participant m, and run n. The following roll-up average over runs, objects, participants, and time then simply excludes all measurements which are either not of interest or not defined:

$$V = \frac{\sum_{k=1}^{K}\sum_{m=1}^{M}\sum_{i=1}^{J}\sum_{n=1}^{N} W_{i,m,n,k} V_{i,m,n}(t_k)\Delta t_k}{\sum_{k=1}^{K}\sum_{m=1}^{M}\sum_{i=1}^{J}\sum_{n=1}^{N} W_{i,m,n,k}\Delta t_k},$$

where K is the total number of scoring time intervals, M the number of participants, J the number of objects occurring in the scenario, and N the number of runs. This formulation has the additional mathematical advantage that the four individual averages that it encompasses can be carried out in any order. This feature allows for considerable flexibility in the definitions of intermediate averages on the way to the final roll-up. For example, if an average value V_m specific to participant m is required, one simply omits the summations over m in the above formula. The final roll-up value V can then be re-formulated as a weighted average of the individual V_ms over participants, with the weights being $\sum_{k=1}^{K}\sum_{i=1}^{J}\sum_{n=1}^{N} W_{i,m,n,k}\Delta t_k$. The flexibility as to intermediate averages may prove useful, not only for various supplementary analysis purposes, but also in adapting the attribute measures to support differing interpretations of CRD requirements (until

agreement is reached) or changes to the requirements themselves in future CRD updates.

Explicit roll-up averages are not provided for the ballistic missile SIAP MOPs, on the grounds that the MOPs may be put to a variety of uses in SIAP system engineering analyses, and that the particular use should determine the appropriate roll-up (if any). There is presently no anticipated use of roll-up MOPs for requirements purposes.

Completeness: Object Track Completeness

Definition: Proportion of real objects meeting reporting criteria that are held as declared tracks at each scheduled scoring time in the scenario.

Calculation: First, determine the set of real objects at time t, during run n that meet reporting criteria in run n, and the cardinality of that set, $N_{ShouldTrk,n}(t)$. Next, for each run at each scheduled track scoring time $t_{ScoreTrk_k}$, identify a gated unique assignment of declared tracks to the set of real object representations according to the invoked metrics tracks-to-truth assignment method. This will yield the following sets/quantities: (1) valid tracks held by participant m during run n, of quantity $NVT_{m,n}(t_{ScoreTrk_k})$, on truth objects meeting reporting criteria; (2) valid "spatial map" tracks on other truth object representations (Method 2 only); (3) extra tracks not assigned to any truth object representations; and (4) missed reportable objects. The instantaneous object track completeness for participant m during run n at the k^{th} track scoring time is then:

$$\text{Track\_Completeness}_{m,n}(t_{ScoreTrk_k}) = \frac{NVT_{m,n}(t_{ScoreTrk_k})}{N_{ShouldTrk,n}(t_{ScoreTrk_k})}.$$

The Object Track Completeness may be averaged over runs (in a reportable-object-weighted sense) for the particular participant m at the k^{th} track scoring time, as follows:

$$\text{Track\_Completeness}_{m}(t_{SoreTrk_K}) = \frac{\sum_{n=1}^{N} NVT_{m,n}(t_{ScoreTrk_K})}{\sum_{n=1}^{N} N_{ShouldTrk,n}(t_{ScoreTrk_K})},$$

where N is the number of runs.

An instantaneous (weighted) average of Object Track Completeness over runs and participants at the track scoring time $t_{ScoreTrk_k}$ may be obtained as follows:

$$\text{Track\_Completeness}(t_{ScoreTrk_K}) = \frac{\sum_{m=1}^{M}\sum_{n=1}^{N} NVT_{m,n}(t_{ScoreTrk_K})}{\sum_{m=1}^{M}\sum_{n=1}^{N} N_{ShouldTrk,n}(t_{ScoreTrk_K})},$$

where M is the number of participants.

The Object Track Completeness may then be averaged over time (in a reportable-object-weighted sense) where the summations are over all track scoring times. The roll-up average over time is given by:

$$\text{Track\_Completeness} = \frac{\sum_{k=1}^{K}\sum_{m=1}^{M}\sum_{n=1}^{N} NVT_{m,n}(t_{ScoreTrk_K})\Delta t_{ScoreTrk_K}}{\sum_{k=1}^{K}\sum_{m=1}^{M}\sum_{n=1}^{N} N_{ShouldTrk,n}(t_{ScoreTrk_K})\Delta t_{ScoreTrk_K}},$$

where K is the total number of scoring time subintervals and $\Delta t_{ScoreTrk_K}$ is the duration of the k^{th} time interval.

Completeness: Launch Point Completeness

Definition: Proportion of real ballistic missiles that have been launched for which a valid launch point estimate is being provided.

Calculation: First, determine the set of real ballistic missile launches that have occurred by time t in the scenario and the size (number) of that set, $N_{Launches,n}(t)$.

At each scheduled LPE scoring time $t_{ScoreLPE_k}$ of each run, obtain from the participating unit m the set of declared LPEs.

For the scheduled LPE scoring of that run, identify a gated unique optimal assignment of declared launch time/point estimates to the set of real launches of ballistic missiles that have occurred by that time into the scenario during run n. (Follow the assignment procedure delineated in 3.3.4.1 and 3.3.4.2.) The gated unique optimal assignment will yield three sets and corresponding quantities: (1) valid assigned LPEs, of quantity $NVL_{m,n}(t_{ScoreLPE_k})$; (2) extra LPEs; and (3) missed launch points. The instantaneous LPE completeness for participant m during run n is then

$$\text{LPE\_Completeness}_{m,n}(t_{ScoreLPE_k}) = \frac{NVL_{m,n}(t_{ScoreLPE_k})}{N_{Launches,n}(t_{ScoreLPE_k})}.$$

The average LPE Completeness over runs for participant m at the k^{th} LPE scoring time is computed in a reportable-object-weighted sense, as follows:

$$LPE\_Completeness_m(t_{ScoreLPE_k}) = \frac{\sum_{n=1}^{N} NVL_{m,n}(t_{ScoreLPE_k})}{\sum_{n=1}^{N} N_{Launches,n}(t_{ScoreLPE_k})}.$$

Similarly, the instantaneous (weighted) average over runs and participants is:

$$LPE\_Completeness(t_{ScoreLPE\,k}) = \frac{\sum_{m=1}^{M}\sum_{n=1}^{N} NVL_{m,n}(t_{ScoreLPE_k})}{\sum_{m=1}^{M}\sum_{n=1}^{N} N_{Launches,n}(t_{ScoreLPE_k})}.$$

The roll-up LPE Completeness over runs, participants, and time is defined as the ratio of the maxima of the numerator and denominator in the instantaneous average formula:

$$LPE\_Completeness = \frac{\max_{k\in[1,K]} \sum_{m=1}^{M}\sum_{n=1}^{N} NVL_{m,n}(t_{ScoreLPE_k})}{\max_{k\in[1,K]} \sum_{m=1}^{M}\sum_{n=1}^{N} N_{Launches,n}(t_{ScoreLPE_k})},$$

where K is the total number of scoring time subintervals. Note that the denominator in this formula, since it is monotonically nondecreasing, is simply equal to the value of $\sum_{m=1}^{M}\sum_{n=1}^{N} N_{Launches,n}(t) = M\sum_{n=1}^{N} N_{Launches,n}(t)$ at the final scoring time.

Completeness: Impact Point Completeness

Definition: Proportion of objects meeting reporting criteria that have a valid impact point prediction.

Calculation: First, determine the time-dependent set of real objects that meet reporting criteria in run n, and the quantity, $N_{Reportable,n}(t)$, in that set. Cross reference their object numbers to ground impact of truth trajectories having (successive) primary-child inheritance, i.e., impacts of reportable primary-child threads. Assign a new index $j \in [1,P]$ to each such reportable object thread, where P is the total number of reportable threads.

At each scheduled IPP scoring time $t_{ScoreIPP_k}$ of each run, each participating unit m should provide its set of declared IPPs.

For the k^{th} scheduled IPP scoring of run n, identify a gated unique assignment of declared IPPs to the set of true currently reportable object thread impacts. (Follow the procedure delineated in 3.3.4.1 and 3.3.4.3.) These procedures will yield three sets and corresponding quantities: (1) valid assigned IPPs held by participant m, of quantity $NVI_{m,n}(t_{ScoreIPP_k})$, (2) extra IPPs, and (3) missed reportable object thread impacts.

If the intent is to score IPP Completeness only for particular intervals of time[1] $I_{j,n}$ prior to the impact of each reportable object, then define the Boolean variable $WI_{j,n,k}$ to be 1 if $t_{ScoreIPP_k} \in I_{j,n}$ and object thread j is currently reportable at $t_{ScoreIPP_k}$. Otherwise, set $WI_{j,n,k}$ to 0. If no such intervals are specified, then take $I_{j,n}$ to be the entire time period of the scenario, making $N_{Reportable,n}(t_{ScoreIPP_k}) = \sum_{j=1}^{P} WI_{j,n,k}$. Also set $WCI_{j,m,n,k}$ to 1 if object thread j is currently reportable and its impact is covered by a valid IPP held by participant m at scoring time $t_{ScoreIPP_k}$; otherwise set $WCI_{j,m,n,k}$ to 0. Note that $NVI_{m,n}(t_{ScoreIPP_k}) = \sum_{j=1}^{P} WCI_{j,m,n,k}$, since assigned IPPs are in one-to-one correspondence with covered reportable object thread impacts.

The instantaneous IPP completeness for participant m at the k^{th} scoring time during run n is defined by:

$$IPP\_Completeness_{m,n}(t_{ScoreIPP_k}) = \frac{\sum_{j=1}^{P} WI_{j,m,n,k} WCI_{j,n,k}}{\sum_{j=1}^{P} WI_{j,n,k}}.$$

The average IPP completeness over runs for participant m at $t_{ScoreIPP_k}$ is given by:

$$IPP\_Completeness_{m}(t_{ScoreIPP_k}) = \frac{\sum_{n=1}^{N}\sum_{j=1}^{P} WI_{j,m,n,k} WCI_{j,n,k}}{\sum_{n=1}^{N}\sum_{j=1}^{P} WI_{j,n,k}}.$$

The instantaneous average over runs and participants is

[1] The current TAMD CRD (2001) does not specify any such intervals, but it is believed that this is the intent.

$$\text{IPP\_Completeness}\ (t_{ScoreIPP_k}) = \frac{\sum_{m=1}^{M}\sum_{n=1}^{N}\sum_{j=1}^{P} WI_{j,m,n,k}\, WCI_{j,n,k}}{\sum_{m=1}^{M}\sum_{n=1}^{N}\sum_{j=1}^{P} WI_{j,n,k}},$$

and the roll-up over runs, participants, and time is

$$\text{IPP\_Completeness} = \frac{\sum_{k=1}^{K}\sum_{m=1}^{M}\sum_{n=1}^{N}\sum_{j=1}^{P} WI_{j,m,n,k}\, WCI_{j,n,k}\, \Delta t_{ScoreIPP_k}}{\sum_{k=1}^{K}\sum_{m=1}^{M}\sum_{n=1}^{N}\sum_{j=1}^{P} WCI_{j,n,k}\, \Delta t_{ScoreIPP_k}},$$

where K is the number of scoring time intervals and $\Delta t_{ScoreIPP_k}$ is the duration of the k^{th} interval.

Clarity: Track Ambiguity

Definition: In a gated non-unique assignment using Method 1, the number of declared tracks that are assignable to real objects meeting reporting criteria, divided by the number of declared tracks are actually assigned to reportable objects. (Track ambiguity is currently undefined for truth-to-tracks assignment Method 2.)

Calculation: For each run, each participant m, and each scheduled track scoring time, compute the Method 1 assignable declared tracks, i.e., the declared tracks that are within the gates of real objects meeting reporting criteria. This will yield the number $NAT_{m,n}(t_{ScoreTrk_k})$ of assignable declared tracks.

From the assignable tracks, using Method 1 perform a unique assignment of declared tracks to the set of real objects meeting reporting criteria. This will identify the valid tracks, of quantity $NVT_{m,n}(t_{ScoreTrk_k})$, on truth objects meeting reporting criteria.

Track ambiguity for participant m, for a single run n at the k^{th} scheduled track scoring time is given by:

$$RTR_{m,n}(t_{ScoreTrk_k}) = \frac{NAT_{m,n}(t_{ScoreTrk_k})}{NVT_{m,n}(t_{ScoreTrk_k})}.$$

The (valid track, or tracked object, weighted) average of track ambiguity over runs for participant m at the k^{th} scoring time is then:

$$RTR_m(t_{ScoreTrk_k}) = \frac{\sum_{n=1}^{N} NAT_{m,n}(t_{ScoreTrk_k})}{\sum_{n=1}^{N} NVT_{m,n}(t_{ScoreTrk_k})}.$$

Track ambiguity can then be instantaneously averaged over participants at $t_{ScoreTrk_k}$ as follows:

$$RTR(t_{ScoreTrk_k}) = \frac{\sum_{m=1}^{M}\sum_{n=1}^{N} NAT_{m,n}(t_{ScoreTrk_k})}{\sum_{m=1}^{M}\sum_{n=1}^{N} NVT_{m,n}(t_{ScoreTrk_k})}$$

Finally, the roll-up average over time of track ambiguity is given by:

$$RTR = \frac{\sum_{k=1}^{K}\sum_{m=1}^{M}\sum_{n=1}^{N} NAT_{m,n}(t_{ScoreTrk_k})\Delta t_{ScoreTrk_k}}{\sum_{k=1}^{K}\sum_{m=1}^{M}\sum_{n=1}^{N} NVT_{m,n}(t_{ScoreTrk_k})\Delta t_{ScoreTrk_k}}.$$

Clarity: Spurious Track Mean Ratio

Definition: In a gated non-unique assignment using Method 2, the number of declared tracks that are unassignable to any real object, divided by the number of declared tracks. (Spurious Track Mean Ratio is currently undefined for truth-to-tracks assignment Method 1.)

Calculation: For each run, each participant m, and each scheduled track scoring time, compute the Method 2 unassignable declared tracks, i.e., the declared tracks that are not within the gates of any real object representation. This will yield the number $NUAT_{m,n}(t_{ScoreTrk_k})$ of unassignable declared tracks. Note the number of declared tracks $NT_{m,n}(t_{ScoreTrk_k})$ in the Method 2 assignment of tracks to truth at scheduled scoring time $t_{ScoreTrk_k}$.

The spurious track ratio for participant m, for a single run n at the kth scheduled track scoring time is given by:

$$STR_{m.n}(t_{ScoreTrk_k}) = \frac{NUAT_{m,n}(t_{ScoreTrk_k})}{NT_{m,n}(t_{ScoreTrk_k})}.$$

The (declared track weighted) instantaneous average of the spurious track ratio over runs for participant m is given by:

$$STR_m(t_{ScoreTrk_k}) = \frac{\sum_{n=1}^{N} NUAT_{m,n}(t_{ScoreTrk_k})}{\sum_{n=1}^{N} NT_{m,n}(t_{ScoreTrk_k})}.$$

The spurious track ratio can then be averaged over participants at scoring time $t_{ScoreTrk_k}$ as follows:

$$STR(t_{ScoreTrk_k}) = \frac{\sum_{m=1}^{M}\sum_{n=1}^{N} NUAT_{m,n}(t_{ScoreTrk_k})}{\sum_{m=1}^{M}\sum_{n=1}^{N} NT_{m,n}(t_{ScoreTrk_k})}$$

Finally, the roll-up average over time of the spurious track ratio, where the summations are over all track scoring times, is given by:

$$STR = \frac{\sum_{k=1}^{K}\sum_{m=1}^{M}\sum_{n=1}^{N} NUAT_{m,n}(t_{ScoreTrk_k})\Delta t_{ScoreTrk_k}}{\sum_{k=1}^{K}\sum_{m=1}^{M}\sum_{n=1}^{N} NT_{m,n}(t_{ScoreTrk_k})\Delta t_{ScoreTrk_k}}.$$

Clarity: LPE Ambiguity

Definition: In a gated non-unique assignment, the number of declared launch point/time estimates that are assignable to real ballistic missile launches, divided by the number of valid declared LPEs.

Calculation: Use the set of real ballistic missile launches that have occurred by time t in the scenario and the quantity $NVL_{m,n}(t_{ScoreLPE_k})$ from the calculation of launch point completeness.

For each run, each participant m, and each scheduled LPE scoring time, compute the assignable declared LPEs, i.e., the declared LPEs that are within gates with real launches of ballistic missiles that have occurred by that time in the scenario. (Follow the LPE gating procedures described for computing LPE completeness.) This will yield two groupings and corresponding counts for each scheduled LPE scoring time $t_{ScoreLPE_k}$: (1) number $NAL_{m,n}(t_{ScoreLPE_k})$ of assignable declared LPEs and (2) number $NUAL_{m,n}(t_{ScoreLPE_k})$ of unassignable declared LPEs, the latter being precluded by the gating constraint.

LPE ambiguity for participant m, for a single run n at the kth scheduled LPE scoring time is given by:

$$RLR_{m,n}(t_{ScoreLPE_k}) = \frac{NAL_{m,n}(t_{ScoreLPE_k})}{NVL_{m,n}(t_{ScoreLPE_k})}.$$

The (valid LPE, or covered launch, weighted) average of LPE ambiguity over runs for participant m at the kth LPE scoring time is given by:

$$RLR_m(t_{ScoreLPE_k}) = \frac{\sum_{n=1}^{N} NAL_{m,n}(t_{ScoreLPE_k})}{\sum_{n=1}^{N} NVL_{m,n}(t_{ScoreLPE_k})}$$

LPE ambiguity can then be averaged over participants at the kth LPE scoring time as follows:

$$RLR(t_{ScoreLPE_k}) = \frac{\sum_{m=1}^{M} \sum_{n=1}^{N} NAL_{m,n}(t_{ScoreLPE_k})}{\sum_{m=1}^{M} \sum_{n=1}^{N} NVL_{m,n}(t_{ScoreLPE_k})}$$

Finally, the roll-up average of LPE ambiguity over time, where the summations are over all track scoring times, is given by:

$$RLR = \frac{\sum_{k=1}^{K} \sum_{m=1}^{M} \sum_{n=1}^{N} NAL_{m,n}(t_{ScoreLPE_k})\Delta t_{ScoreLPE_k}}{\sum_{k=1}^{K} \sum_{m=1}^{M} \sum_{n=1}^{N} NVL_{m,n}(t_{ScoreLPE_k})\Delta t_{ScoreLPE_k}}.$$

Clarity: Spurious LPE Mean Ratio

Definition: In a gated non-unique assignment, the number of declared launch point/time estimates that are unassignable to real ballistic missile launches, divided by the number of declared distinct LPEs.

Calculation: Use the number $NUAL_{m,n}(t_{ScoreLPE_k})$ of unassignable declared LPEs obtained in computing LPE ambiguity.

For a run n and the kth scheduled LPE scoring time, the spurious LPE ratio for participant m is given by:

$$SLR_{m,n}(t_{ScoreLPE_k}) = \frac{NUAL_{m,n}(t_{ScoreLPE_k})}{NDL_{m,n}(t_{ScoreLPE_k})},$$

where $NDL_{m,n}(t_{ScoreLPE_k})$ is the number LPEs held and declared by participant m during run n at scoring time $t_{ScoreLPE_k}$.

The (declared LPE weighted) averaged of the spurious LPE ratio over runs for participant m at the k^{th} LPE scoring time is then:

$$SLR_m(t_{ScoreLPE_k}) = \frac{\sum_{n=1}^{N} NUAL_{m,n}(t_{ScoreLPE_k})}{\sum_{n=1}^{N} NDL_{m,n}(t_{ScoreLPE_k})}$$

Then the spurious LPE ratio can be averaged over participants at the k^{th} scoring time as follows:

$$SLR(t_{ScoreLPE_k}) = \frac{\sum_{m=1}^{M}\sum_{n=1}^{N} NUAL_{m,n}(t_{ScoreLPE_k})}{\sum_{m=1}^{M}\sum_{n=1}^{N} NDL_{m,n}(t_{ScoreLPE_k})}.$$

Finally, the roll-up average of spurious LPE ratio over time, where the summations are over all LPE scoring times, is given by:

$$SLR = \frac{\sum_{k=1}^{K}\sum_{m=1}^{M}\sum_{n=1}^{N} NUAL_{m,n}(t_{ScoreLPE_k})\Delta t_{ScoreLPE_k}}{\sum_{k=1}^{K}\sum_{m=1}^{M}\sum_{n=1}^{N} NDL_{m,n}(t_{ScoreLPE_k})\Delta t_{ScoreLPE_k}}.$$

Clarity: IPP Ambiguity

Definition: In a gated non-unique assignment, the number of declared impact point/time predictions that are assignable to true impact points/times of currently reportable object threads, divided by the number of valid declared IPPs.

Calculation: Use the set of true currently reportable object thread impacts and the quantity $NVI_{m,n}(t_{ScoreIPP_k})$ of valid IPPs from the calculation of IPP completeness.

For each run, each participant m, and each scheduled IPP scoring time, compute the assignable declared IPPs, i.e., the declared IPPs that are within gates with true currently reportable object thread impact points/times. (Follow the IPP gating procedures described for computing IPP completeness.) This will yield two groupings and corresponding counts for each scheduled IPP scoring time $t_{ScoreIPP_k}$: (1) number $NAI_{m,n}(t_{ScoreIPP_k})$ of assignable declared IPPs and (2) number $NUAI_{m,n}(t_{ScoreIPP_k})$ of unassignable declared IPPs, the latter being precluded by the gating constraint.

IPP ambiguity for participant m, for a single run N at the kth scheduled IPP scoring time is given by:

$$RIR_{m,n}(t_{ScoreIPP_k}) = \frac{NAI_{m,n}(t_{ScoreIPP_k})}{NVI_{m,n}(t_{ScoreIPPk})}$$

The (valid IPP, or covered impact, weighted) averaged of IPP ambiguity over runs for participant m at the kth IPP scoring time is then:

$$RIR_m(t_{ScoreIPP_k}) = \frac{\sum_{n=1}^{N} NAI_{m,n}(t_{ScoreIPP_k})}{\sum_{n=1}^{N} NVI_{m,n}(t_{ScoreIPPk})}.$$

Then IPP ambiguity can be averaged over participants at the kth scoring time as follows:

$$RIR(t_{ScoreIPP_k}) = \frac{\sum_{m=1}^{M}\sum_{n=1}^{N} NAI_{m,n}(t_{ScoreIPP_k})}{\sum_{m=1}^{M}\sum_{n=1}^{N} NVI_{m,n}(t_{ScoreIPPk})}$$

Finally, the roll-up average of IPP ambiguity over time, where the summations are over all IPP scoring times, is given by:

$$RIR = \frac{\sum_{k=1}^{K}\sum_{m=1}^{M}\sum_{n=1}^{N} NAI_{m,n}(t_{ScoreIPP_k})\Delta t_{ScoreIPP_k}}{\sum_{k=1}^{K}\sum_{m=1}^{M}\sum_{n=1}^{N} NVI_{m,n}(t_{ScoreIPPk})\Delta t_{ScoreIPP_k}}.$$

Clarity: Spurious IPP Mean Ratio

Definition: In a gated non-unique assignment, the number of declared impact point/time predictions that are unassignable to true impact points/times of currently reportable object threads, divided by the number of declared distinct IPPs.

Calculation: Use the number $NUAI_{m,n}(t_{ScoreIPP_k})$ of unassignable declared IPPs obtained in computing IPP ambiguity.

For a run n and the k[th] scheduled IPP scoring time, the spurious IPP ratio for participant m is given by:

$$SIR_{m,n}(t_{ScoreIPP_k}) = \frac{NUAI_{m,n}(t_{ScoreIPP_k})}{NDI_{m,n}(t_{ScoreIPP_k})},$$

where $NDI_{m,n}(t_{ScoreIPP_k})$ is the number of IPPs held and declared by participant m during run n at scoring time $t_{ScoreIPP_k}$.

The (declared IPP weighted) average of spurious IPP ratio over runs for participant m at the k[th] IPP scoring time is then:

$$SIR_m(t_{ScoreIPP_k}) = \frac{\sum_{n=1}^{N} NUAI_{m,n}(t_{ScoreIPP_k})}{\sum_{n=1}^{N} NDI_{m,n}(t_{ScoreIPP_k})}.$$

Then the spurious IPP ratio can be averaged over participants at the k[th] scoring time as follows:

$$SIR(t_{ScoreIPP_k}) = \frac{\sum_{m=1}^{M}\sum_{n=1}^{N} NUAI_{m,n}(t_{ScoreIPP_k})}{\sum_{m=1}^{M}\sum_{n=1}^{N} NDI_{m,n}(t_{ScoreIPP_k})}.$$

Finally, the roll-up average of spurious IPP ratio over time, where the summations are over all IPP scoring times, is given by:

$$SIR = \frac{\sum_{k=1}^{K}\sum_{m=1}^{M}\sum_{n=1}^{N} NUAI_{m,n}(t_{ScoreIPP_k})\Delta t_{ScoreIPP_k}}{\sum_{k=1}^{K}\sum_{m=1}^{M}\sum_{n=1}^{N} NDI_{m,n}(t_{ScoreIPP_k})\Delta t_{ScoreIPP_k}}.$$

Continuity: Normalized Longest Duration Valid Track Segment on Distinct Reportable Objects

Definition: For a distinct real object meeting reporting criteria, the longest duration valid track segment without a track number switch or track break, normalized by the scenario lifetime of the object.

Calculation: Use the determination of the time-dependent set of real objects meeting reporting criteria from the calculation of track completeness. Determine the lifetime of each such object i during run n as

$LT_{i,n}$ = min[(scenario time), (time duration of reportable object i in run n)].

Let $LVT_{i,m,n}$ denote the longest duration (up to the current time) of a valid track held by participant m on reportable object i during run n. At the beginning of each run n, initialize all $LVT_{i,m,n} = 0$ and all $VT_{i,m,n} = 0$. (The variable $VT_{i,m,n}$ will serve as a time counter for the candidate track currently being assessed.)

At each scheduled track scoring time $t_{ScoreTrk}$ of each run n, identify the gated unique assignment of declared tracks to the set of real object representations, according to the invoked metrics assignment method. If (1) a declared track is assigned to reportable object i this scoring time and (2) the track number assigned to object i this scheduled scoring time was also assigned last scheduled track scoring time to object i, then increment $VT_{i,m,n}$ by that scoring time interval. Otherwise, reset $VT_{i,m,n} = 0$. If $VT_{i,m,n}$ is greater than $LVT_{i,m,n}$ then set $LVT_{i,m,n} = VT_{i,m,n}$.

For each participant m and each reportable object i of interest, histogram the statistical density and plot the cumulative distribution of quantities ($LVT_{i,m,n}$ / $LT_{i,n}$) for the results from the set of runs.

The Normalized Longest Duration Valid Track Segment for participant m on reportable object i is computed as an object lifetime-weighted average over runs:

$$NLVT_{i,m} = \frac{\sum_{n=1}^{N} LVT_{i,m,n}}{\sum_{n=1}^{N} LT_{i,n}},$$

The overall measure for a given participant m is then computed as an analogously weighted average over reportable objects:

$$NLVT_m = \frac{\sum_{i=1}^{L}\sum_{n=1}^{N} LVT_{i,m,n}}{\sum_{i=1}^{L}\sum_{n=1}^{N} LT_{i,n}},$$

where L is the number of reportable objects over the duration of the scenario. Finally, the roll-up average of the Normalized Longest Duration Valid Track Segment on Reportable Objects is given by:

$$NLVT = \frac{\sum_{m=1}^{M}\sum_{i=1}^{L}\sum_{n=1}^{N} LVT_{i,m,n}}{M\sum_{i=1}^{L}\sum_{n=1}^{N} LT_{i,n}}.$$

Continuity: Normalized Longest Duration Valid Track Segment on Reportable Object Threads

Definition: For a real primary-child thread of objects meeting reporting criteria, the longest duration valid track segment without a track numbers switch or track break, normalized by the scenario lifetime of the object thread.

Calculation: Using the determination of the time-dependent set of real objects meeting reporting criteria from the calculation of completeness, determine the corresponding set of reportable parent-primary child object threads and the lifetime of each such object thread j as

$LTT_{j,n}$ = min[(scenario time), (time duration of object thread j containing a reportable object in run n)]

Let $LVTT_{j,m,n}$ denote the longest duration (up to the current time) of a valid track held by participant m on reportable object thread j during run n. At the beginning of each run n, initialize all $LVTT_{j,m,n} = 0$ and all $VTT_{j,m,n} = 0$. (The variable $VTT_{j,m,n}$ will serve as a time counter for the candidate track currently being assessed.)

At each scheduled track scoring time $t_{ScoreTrk}$ of each run n, identify the gated unique assignment of declared tracks to the set of real object representations according to the invoked metrics assignment method. If (1) object thread j contains reportable object i and (2) a declared track is assigned to object i this scoring time and (3) the track number assigned to object i this scheduled scoring time was also assigned last scheduled track scoring time to object i or to a parent/grandparent object in thread j, then increment $VTT_{j,m,n}$ by that scoring time interval. Otherwise, reset $VTT_{j,m,n} = 0$. If $VTT_{j,m,n}$ is greater than $LVTT_{j,m,n}$ then set $LVTT_{j,m,n} = VTT_{j,m,n}$.

For each participant m and each reportable object thread j of interest, histogram the statistical density and plot the cumulative distribution of quantities ($LVTT_{j,m,n}$ / LTT_j) for the results from the set of runs.

The Normalized Longest Duration Valid Track Segment for participant m on reportable object thread j is computed as an object thread lifetime-weighted average over runs:

$$NLVTT_{j,m} = \frac{\sum_{n=1}^{N} LVTT_{j,m,n}}{\sum_{n=1}^{N} LTT_{j,n}},$$

The overall measure for a given participant m is then computed as an analogously weighted average over reportable object threads:

$$NLVTT_{m} = \frac{\sum_{j=1}^{P}\sum_{n=1}^{N} LVTT_{j,m,n}}{\sum_{j=1}^{P}\sum_{n=1}^{N} LTT_{j,n}}.$$

Finally, the roll-up average of the Normalized Longest Duration Valid Track Segment on Reportable Object Threads is given by:

$$NLVTT = \frac{\sum_{m=1}^{M}\sum_{j=1}^{P}\sum_{n=1}^{N} LVTT_{j,m,n}}{M\sum_{j=1}^{P}\sum_{n=1}^{N} LTT_{j,n}}.$$

Track Continuity: Cumulative Switches of Tracks on Distinct Reportable Objects

Definition: For distinct real objects meeting reporting criteria, the cumulative number of switches of tracks (not counting breaks) by time t into the scenario for particular objects, and summed across all reportable objects.

Calculation: Use the determination of the time-dependent set of real objects meeting reporting criteria from the calculation of track completeness. Initialize all $NS_{i,m,n}(t_{ScoreTrk}) = 0$.

For each participant m, at each scheduled track scoring time $t_{ScoreTrk}$ of each run n, identify the gated unique assignment of declared tracks to the set of real object representations, according to the invoked metrics assignment method.

For each object i meeting reporting criteria, if track j held by participant m was assigned to object i at the last scheduled track scoring time of this run and track k (k not equal j) is assigned to object i at the current track scoring time $t_{ScoreTrk}$, then increment by one the number of switches $NS_{i,m,n}(t_{ScoreTrk})$ of track on object i.

Average across all runs for each object i as $NS_{i,m}(t_{ScoreTrk}) = \frac{1}{N}\sum_{n=1}^{N} NS_{i,m,n}(t_{ScoreTrk})$, where N is the number of runs, and plot $NS_{i,m}(t)$ versus time for particular objects of interest.

Sum $NS_{i,m}(t_{ScoreTrk})$ across all reportable objects in the scenario as $SNS_m(t_{ScoreTrk}) = \sum_{i=1}^{L} NS_{i,m}(t_{ScoreTrk})$, and plot this sum $SNS_m(t)$ versus time for each participant m.

Higher-level roll-ups of this MOP may be computed as required by the user.

Track Continuity: Cumulative Switches of Tracks on Reportable Object Threads

Definition: For real primary-child threads of objects meeting reporting criteria, the cumulative number of switches of tracks (not counting breaks) by time t into the scenario for particular threads and summed across all reportable object threads.

Calculation: Use the determination of the time-dependent set of real objects meeting reporting criteria from the calculation of completeness. Determine the history of acknowledged parent truth object(s) for every real object i meeting reporting criteria. Initialize all $NS_{j,m,n}(t_{ScoreTrk}) = 0$.

For each participant m, at each scheduled track scoring time $t_{ScoreTrk}$ of each run n, identify the gated unique assignment of declared tracks to the set of real object representations according to the invoked metrics assignment method. For each object i meeting reporting criteria, if track p held by participant m was assigned to object i or to one of the parent/grandparent objects in the thread for object i at the last scheduled track scoring time of this run and track q (q not equal to p) is assigned to object i at the current track scoring time $t_{ScoreTrk}$, then increment by one the number of switches $NS_{j,m,n}(t_{ScoreTrk})$ of track on the parent-primary child object thread j containing object i.

Average across all runs for each parent-primary child object thread j as $NS_{j,m}(t_{ScoreTrk}) = \frac{1}{N}\sum_{n=1}^{N} NS_{j,m,n}(t_{ScoreTrk})$ and plot $NS_{j,m}(t)$ versus time for particular object threads of interest.

Sum $NS_{j,m}(t_{ScoreTrk})$ across all reportable parent-primary child object threads in the scenario as $SNS_m(t_{ScoreTrk}) = \sum_{j=1}^{P} NS_{j,m}(t_{ScoreTrk})$, and plot this sum $SNS_m(t)$ versus time for each participant.

Higher-level roll-ups of this MOP may be computed as required by the user.

Track Continuity: Cumulative Broken Tracks on Distinct Reportable Objects

Definition: For distinct real objects meeting reporting criteria, the cumulative number of breaks of tracks by time t into the scenario for particular objects, and summed across all reportable objects.

Calculation: Use the determination of the time-dependent set of real objects meeting reporting criteria from the calculation of track completeness. Initialize all $NB_{i,m,n}(t_{ScoreTrk}) = 0$.

For each participant m, at each scheduled track scoring time $t_{ScoreTrk}$ of each run n, identify the gated unique assignment of declared tracks to the set of real object representations according to the invoked metrics assignment method. For each object i meeting reporting criteria, if a track held by participant m was assigned to object i at the last scheduled track scoring time of this run and no track is assigned to object i at the current track scoring time $t_{ScoreTrk}$ (and object i still exists), then increment by one the number of breaks $NB_{i,m,n}(t_{ScoreTrk})$ of track on object i.

Average across all runs for each object i as $NB_{i,m}(t_{ScoreTrk}) = \frac{1}{N}\sum_{n=1}^{N} NB_{i,m,n}(t_{ScoreTrk})$ and plot $NB_{i,m}(t)$ versus time for particular objects of interest.

Sum $NB_{i,m}(t_{ScoreTrk})$ across all reportable objects in the scenario as $SNB_m(t_{ScoreTrk}) = \sum_{i=1}^{L} NB_{i,m}(t_{ScoreTrk})$ and plot this sum $SNB_m(t)$ versus time for each participant.

Higher-level roll-ups of this MOP may be computed as required by the user.

Track Continuity: Cumulative Broken Tracks on Reportable Object Threads

Definition: For real primary-child threads of objects meeting reporting criteria, the cumulative number of breaks of tracks by time t into the scenario for particular threads, and summed across all reportable object threads.

Calculation: Use the determination of the time-dependent set of real objects meeting reporting criteria from the calculation of track completeness. Determine the history of parent truth object(s) for every real object i meeting reporting criteria. Initialize all $NB_{j,m,n}(t_{ScoreTrk}) = 0$.

For each participant m, at each scheduled track scoring time $t_{ScoreTrk}$ of each run n, identify the gated unique assignment of declared tracks to the set of real object representations according to the invoked metrics assignment method. For each object i meeting reporting criteria, if a track held by participant m was assigned to object i or to one of the parent/grandparent objects in the thread for object i at the last scheduled track scoring time of this run and no track is assigned to object i at the current track scoring time $t_{ScoreTrk}$, then increment by one the number of breaks $NB_{j,m,n}(t_{ScoreTrk})$ of track on the parent-primary child object thread j containing object i.

Average across all runs for each parent-primary child object thread j as $NB_{j,m}(t_{ScoreTrk}) = \frac{1}{N}\sum_{m=1}^{N} NB_{j,m,n}(t_{ScoreTrk})$ and plot $NB_{j,m}(t)$ versus time for particular object threads of interest.

Sum $NB_{j,m}(t_{ScoreTrk})$ across all reportable parent-primary child object threads in the scenario as $SNB_m(t_{ScoreTrk}) = \sum_{j=1}^{P} NB_{j,m}(t_{ScoreTrk})$ and plot this sum $SNB_m(t)$ versus time for each participant.

Higher-level roll-ups of this MOP may be computed as required by the user.

Accuracy: Track Accuracy

Definition: As assessed for a particular object meeting reporting criteria, the root mean squared error (RMSE) history in position, the RMSE history in velocity, the root sum squared average error (RSSAE) history in position, and the RSSAE history in velocity of the track assigned to that object compared to the truth states for that object.

Remark: In this and in subsequent calculations employing running averages, the variable n(t) is a counter of the number of runs in which the metric of interest has been updated, as of scoring time t.

Calculation: Determine the set of real objects meeting reporting criteria. Initialize $n(t_{ScoreTrk})$ to 0.

For each participant m, at each scheduled track scoring time $t_{ScoreTrk_k}$ of each run, identify the gated unique assignment of declared tracks to real objects meeting reporting criteria.

The track accuracy attribute measures are defined for each reportable object on which a participant holds track at a particular scoring time. The calculation assumes that these object-specific metrics will be scored only during certain intervals of interest during boost phase and post-boost phase (derived from the TAMD CRD or specified by the user). The object-specific track accuracy measures for reportable object i, participant m, and scoring time $t_{ScoreTrk_k}$ are defined as appropriate averages over runs for which the scoring time is within an interval of interest:

$$RSSAE_{Position,i,m}(t_{ScoreTrk_k}) = \frac{\sum_{n=1}^{N} WA_{i,m,n,k} \left\| \mathbf{e}_{Position,i,m,n}(t_{ScoreTrk_k}) \right\|}{\sum_{n=1}^{N} WA_{i,m,n,k}}$$

$$RSSAE_{Velocity,i,m}(t_{ScoreTrk_k}) = \frac{\sum_{n=1}^{N} WA_{i,m,n,k} \left\| \mathbf{e}_{Velocity,i,m,n}(t_{ScoreTrk_k}) \right\|}{\sum_{n=1}^{N} WA_{i,m,n,k}}$$

$$RMSE_{position,i,m}(t_{ScoreTrk_k}) = \sqrt{\frac{\sum_{n=1}^{N} WA_{i,m,n,k} \left\| \mathbf{e}_{Position,j,m,n}(t_{ScoreTrk_k}) \right\|^2}{\sum_{n=1}^{N} WA_{i,m,n,k}}}$$

$$RMSE_{Velocity,i,m}(t_{ScoreTrk_k}) = \sqrt{\frac{\sum_{n=1}^{N} WA_{i,m,n,k} \left\| \mathbf{e}_{Velocity,j,m,n}(t_{ScoreTrk_k}) \right\|^2}{\sum_{n=1}^{N} WA_{i,m,n,k}}},$$

where the Boolean variable $WA_{i,m,n,k}$ is set to 1 if $t_{ScoreTrk_k}$ is within a time interval of interest for reportable object i in run n and participant m holds a track assigned to object i at $t_{ScoreTrk_k}$, and set to 0 otherwise, $\mathbf{e}_{Position,i,m,n}(t_{ScoreTrk_k})$ and $\mathbf{e}_{Velocity,i,m,n}(t_{ScoreTrk_k})$ are the vectorial differences in position and velocity respectively between reportable truth object i and the aforesaid assigned track, and $\| \; \|$ designates the norm of the vector. Note that the RSSAE and RMSE statistics coincide for a single run.

The run averages may also be computed by the following recursive procedure, performed at each scoring time for a particular object and participant. (Subscripts identifying object, participant, and scoring time increment are suppressed in the recursion formulae, since the key variables are re-initialized at the beginning of each recursion.)

Recursion. If $t_{ScoreTrk}$ is a time of interest for scoring track accuracy for object i in the run under consideration, and a track held by participant m is assigned to object i at that time, increment $n=n(t_{ScoreTrk})$ (i.e., n gets n+1) for that scheduled track scoring time and perform the following recursion updates:

$$\mathbf{e}_n(t_{ScoreTrk}) = \mathbf{x}_n(t_{ScoreTrk}) - \mathbf{x}_{truth}(t_{ScoreTrk});$$
$$\mathbf{d}_n(t_{ScoreTrk}) = \mathbf{e}_n(t_{ScoreTrk}) - \mathbf{e}_{Aver\_n-1}(t_{ScoreTrk});$$
$$\mathbf{e}_{Aver\_n}(t_{ScoreTrk}) = \mathbf{e}_{Aver\_n-1}(t_{ScoreTrk}) + \frac{\mathbf{d}_n(t_{ScoreTrk})}{n};$$
$$\mathbf{C}_{Aver\_n}(t_{ScoreTrk}) = \frac{n-1}{n}\left[\mathbf{C}_{Aver\_n-1}(t_{ScoreTrk}) + \frac{\mathbf{d}_n(t_{ScoreTrk})\mathbf{d}_n(t_{ScoreTrk})^T}{n}\right]$$

where $\mathbf{x}_n(t_{ScoreTrk})$ is the vector of position/velocity state estimates (6 states) of the track assigned to object i at time $t_{ScoreTrk}$, $\mathbf{x}_{truth}(t_{ScoreTrk})$ is the vector of true states at $t_{ScoreTrk}$, $\mathbf{e}_{Aver\_n}(t_{ScoreTrk})$ is the vector of average errors and $\mathbf{C}_{Aver\_n}(t_{ScoreTrk})$ is the statistical covariance of the errors for the n instances in which a track has been assigned at $t_{ScoreTrk}$, and where $\mathbf{e}_{Aver\_0}(t_{ScoreTrk}) = 0$ and $\mathbf{C}_{Aver\_0}(t_{ScoreTrk}) = 0$. If the conditions for incrementing the index $n(t_{ScoreTrk})$ are not met, then go to the next run without updating any variables.

After completing the above recursion for the k^{th} scoring time for participant m and object i of interest, set $NR_{i,m}(t_{ScoreTrk_k}) = n(t_{ScoreTrk})$, and plot this quantity as a function of scoring time (this represents the number of runs included in the average, and is also equal to $\sum_{n=1}^{N} WA_{i,m,n,k}$).

For $e_{P1}(t_{ScoreTrk})$, $e_{P2}(t_{ScoreTrk})$, and $e_{P3}(t_{ScoreTrk})$ being the three orthogonal position error components in $\mathbf{e}_{Aver\_n}(t_{ScoreTrk})$ and $e_{V1}(t_{ScoreTrk})$, $e_{V2}(t_{ScoreTrk})$, and $e_{V3}(t_{ScoreTrk})$ being the three orthogonal velocity error components in $\mathbf{e}_{Aver\_n}(t_{ScoreTrk})$, the root sum squared average error statistics of interest for object i and participant m are:

$$RSSAE_{Position,i,m}(t_{ScoreTrk_k}) = \sqrt{e_{P1}(t_{ScoreTrk})^2 + e_{P2}(t_{ScoreTrk})^2 + e_{P3}(t_{ScoreTrk})^2}$$

$$RSSAE_{Velocity,i,m}(t_{ScoreTrk_k}) = \sqrt{e_{V1}(t_{ScoreTrk})^2 + e_{V2}(t_{ScoreTrk})^2 + e_{V3}(t_{ScoreTrk})^2}$$

For $C_{P11}(t_{ScoreTrk})$, $C_{P22}(t_{ScoreTrk})$, and $C_{P33}(t_{ScoreTrk})$ being the statistical variances (diagonal terms) of the position error components in $\mathbf{C}_{Aver\_n}(t_{ScoreTrk})$ and $C_{V11}(t_{ScoreTrk})$, $C_{V22}(t_{ScoreTrk})$, and $C_{V33}(t_{ScoreTrk})$ being the statistical variances of the velocity error components in $\mathbf{C}_{Aver\_n}(t_{ScoreTrk})$, the root mean squared error statistics of interest are:

$$RMSE_{Position,i,m}(t_{ScoreTrk_k}) = \sqrt{C_{P11}(t_{ScoreTrk}) + C_{P22}(t_{ScoreTrk}) + C_{P33}(t_{ScoreTrk}) + RSSAE_{Position,i,m}(t_{ScoreTrk_k})^2}$$

$$RMSE_{Velocity,i,m}(t_{ScoreTrk_k}) = \sqrt{C_{V11}(t_{ScoreTrk}) + C_{V22}(t_{ScoreTrk}) + C_{V33}(t_{ScoreTrk}) + RSSAE_{Velocity,i,m}(t_{ScoreTrk_k})^2}$$

Plot $RMSE_{Position,i,m}(t)$, $RMSE_{Velocity,i,m}(t)$, $RSSAE_{Position,i,m}(t)$, and $RSSAE_{Velocity,i,m}(t)$ as a function of time t into the scenario for each participant m, each object i of interest, and each scoring time interval within which $NR_{i,m}(t)$ has a nonzero value.

The instantaneous average track accuracy statistics for participant m are computed as averages over all tracked reportable objects which are of interest for scoring at $t_{ScoreTrk_k}$, as follows:

$$RSSAE_{Position,m}(t_{ScoreTrk_k}) = \frac{\sum_{i=1}^{L} NR_{i,m}(t_{ScoreTrk_k}) RSSAE_{Position,i,m}(t_{ScoreTrk_k})}{\sum_{i=1}^{L} NR_{i,m}(t_{ScoreTrk_k})};$$

$$RSSAE_{Velocity,m}(t_{ScoreTrk_k}) = \frac{\sum_{i=1}^{L} NR_{i,m}(t_{ScoreTrk_k}) RSSAE_{Velocity,i,m}(t_{ScoreTrk_k})}{\sum_{i=1}^{L} NR_{i,m}(t_{ScoreTrk_k})};$$

$$RMSE_{Position,m}(t_{ScoreTrk_k}) = \sqrt{\frac{\sum_{i=1}^{L} NR_{i,m}(t_{ScoreTrk_k}) RMSE_{Position,i,m}(t_{ScoreTrk_k})^2}{\sum_{i=1}^{L} NR_{i,m}(t_{ScoreTrk_k})}};$$

$$RMSE_{Velocity,m}(t_{ScoreTrk_k}) = \sqrt{\frac{\sum_{i=1}^{L} NR_{i,m}(t_{ScoreTrk_k}) RMSE_{Velocity,i,m}(t_{ScoreTrk_k})^2}{\sum_{i=1}^{L} NR_{i,m}(t_{ScoreTrk_k})}}.$$

The weights in these formulae are given by $NR_{i,m}(t_{ScoreTrk_k}) = \sum_{n=1}^{N} WA_{i,m,n,k}$, or else extracted from the recursive procedure defined above.

Similarly, the instantaneous averages over objects and participants are given by:

$$RSSAE_{Position}(t_{ScoreTrk_k}) = \frac{\sum_{m=1}^{M}\sum_{i=1}^{L} NR_{i,m}(t_{ScoreTrk_k}) RSSAE_{Position,i,m}(t_{ScoreTrk_k})}{\sum_{m=1}^{M}\sum_{i=1}^{L} NR_{i,m}(t_{ScoreTrk_k})},$$

$$RSSAE_{Velocity}(t_{ScoreTrk_k}) = \frac{\sum_{m=1}^{M}\sum_{i=1}^{L} NR_{i,m}(t_{ScoreTrk_k}) RSSAE_{Velocity,i,m}(t_{ScoreTrk_k})}{\sum_{m=1}^{M}\sum_{i=1}^{L} NR_{i,m}(t_{ScoreTrk_k})},$$

$$RMSE_{Position}(t_{ScoreTrk_k}) = \sqrt{\frac{\sum_{m=1}^{M}\sum_{i=1}^{L} NR_{i,m}(t_{ScoreTrk_k}) RMSE_{Position,i,m} t_{ScoreTrk_k})^2}{\sum_{m=1}^{M}\sum_{i=1}^{L} NR_{i,m}(t_{ScoreTrk_k})}},$$

$$RMSE_{Velocity}(t_{ScoreTrk_k}) = \sqrt{\frac{\sum_{m=1}^{M}\sum_{i=1}^{L} NR_{i,m}(t_{ScoreTrk_k}) RMSE_{Velocity,i,m} t_{ScoreTrk_k})^2}{\sum_{m=1}^{M}\sum_{i=1}^{L} NR_{i,m}(t_{ScoreTrk_k})}},$$

and the roll-up averages over objects, participants, and time are given by:

$$\mathrm{RSSAE}_{\mathrm{Position}} = \frac{\sum_{k=1}^{K}\sum_{m=1}^{M}\sum_{i=1}^{L} \mathrm{NR}_{i,m}(t_{\mathrm{ScoreTrk}_k})\mathrm{RSSAE}_{\mathrm{Position},i,m}(t_{\mathrm{ScoreTrk}_k})\Delta t_{\mathrm{ScoreTrk}_k}}{\sum_{k=1}^{K}\sum_{m=1}^{M}\sum_{i=1}^{L} \mathrm{NR}_{i,m}(t_{\mathrm{ScoreTrk}_k})\Delta t_{\mathrm{ScoreTrk}_k}},$$

$$\mathrm{RSSAE}_{\mathrm{Velocity}} = \frac{\sum_{k=1}^{K}\sum_{m=1}^{M}\sum_{i=1}^{L} \mathrm{NR}_{i,m}(t_{\mathrm{ScoreTrk}_k})\mathrm{RSSAE}_{\mathrm{Velocity},i,m}(t_{\mathrm{ScoreTrk}_k})\Delta t_{\mathrm{ScoreTrk}_k}}{\sum_{k=1}^{K}\sum_{m=1}^{M}\sum_{i=1}^{L} \mathrm{NR}_{i,m}(t_{\mathrm{ScoreTrk}_k})\Delta t_{\mathrm{ScoreTrk}_k}},$$

$$\mathrm{RMSE}_{\mathrm{Position}} = \sqrt{\frac{\sum_{k=1}^{K}\sum_{m=1}^{M}\sum_{i=1}^{L} \mathrm{NR}_{i,m}(t_{\mathrm{ScoreTrk}_k})\mathrm{RMSE}_{\mathrm{Position},i,m}(t_{\mathrm{ScoreTrk}_k})^2\Delta t_{\mathrm{ScoreTrk}_k}}{\sum_{k=1}^{K}\sum_{m=1}^{M}\sum_{i=1}^{L} \mathrm{NR}_{i,m}(t_{\mathrm{ScoreTrk}_k})\Delta t_{\mathrm{ScoreTrk}_k}}},$$

$$\mathrm{RMSE}_{\mathrm{Velocity}} = \sqrt{\frac{\sum_{k=1}^{K}\sum_{m=1}^{M}\sum_{i=1}^{L} \mathrm{NR}_{i,m}(t_{\mathrm{ScoreTrk}_k})\mathrm{RMSE}_{\mathrm{Velocity},i,m}(t_{\mathrm{ScoreTrk}_k})^2\Delta t_{\mathrm{Scoretrk}_k}}{\sum_{k=1}^{K}\sum_{m=1}^{M}\sum_{i=1}^{L} \mathrm{NR}_{i,m}(t_{\mathrm{ScoreTrk}_k})\Delta t_{\mathrm{ScoreTrk}_k}}}.$$

Accuracy: Track Covariance Consistency

Definition: As assessed for a particular object meeting reporting criteria, the mean normalized six-state Chi-squared statistic of the track assigned to that object.

Calculation: Determine the set of real objects meeting reporting criteria. Initialize $n(t_{ScoreTrk})$ to 0.

For each participant m, at each scheduled track scoring time $t_{ScoreTrk}$ of each run, identify the gated unique assignment of declared tracks to real objects meeting reporting criteria. If a track is assigned to object i at that $t_{ScoreTrk}$, increment n (i.e., n gets n+1) for that scheduled track scoring time and perform the following calculation and recursion:

$$\chi_n^2(t_{\mathrm{ScoreTrk}}) = [\mathbf{x}_n(t_{\mathrm{ScoreTrk}}) - \mathbf{x}_{\mathrm{truth}}(t_{\mathrm{ScoreTrk}})]^T \mathbf{P}_n(t_{\mathrm{ScoreTrk}})^{-1}[\mathbf{x}_n(t_{\mathrm{ScoreTrk}}) - \mathbf{x}_{\mathrm{truth}}(t_{\mathrm{ScoreTrk}})]/N_s$$

$$\chi_{\mathrm{Aver\_n}}^2(t_{\mathrm{ScoreTrk}}) = \chi_{\mathrm{Aver\_n-1}}^2(t_{\mathrm{ScoreTrk}}) + \frac{1}{n}[\chi_n^2(t_{\mathrm{ScoreTrk}}) - \chi_{\mathrm{Aver\_n-1}}^2(t_{\mathrm{ScoreTrk}})]$$

where $\mathbf{x}_n(t_{ScoreTrk})$ is the vector of position/velocity state estimates of the track assigned to object i at time $t_{ScoreTrk}$, $\mathbf{x}_{truth}(t_{ScoreTrk})$ is the vector of true states,

$\mathbf{P}_n(t_{ScoreTrk})$ is the estimated position/velocity error covariance matrix of the track assigned to object i, $N_S = 6$ is the number of states, $\chi^2_n(t_{ScoreTrk})$ is the normalized Chi-squared statistic at $t_{ScoreTrk}$, and $\chi^2_{Aver\_n}(t_{ScoreTrk})$ is the average normalized Chi-squared statistic at $t_{ScoreTrk}$, and where $\chi^2_{Aver\_0}(t_{ScoreTrk}) = 0$.

After completing the above recursion for participant m and object i of interest, plot $n(t_{ScoreTrk})$ as a function of $t_{ScoreTrk}$. (this represents the number of runs for which participant m holds a track assigned to object i at the given scoring time). For time segments where $n(t_{ScoreTrk}) > 0$, set $\chi^2_{i,m}(t_{ScoreTrk}) = \chi^2_{Aver\_n}(t_{ScoreTrk})$.

Plot $\chi^2_{i,m}(t)$ as a function of time t into the scenario for each participant m, each object i of interest, and each scoring time interval within which the statistic has a defined value.

Higher-level roll-ups of this MOP may be computed as required by the user.

Accuracy: LPE Position and Time Accuracy

Definition: As assessed for a particular real ballistic missile launch, the RMSE in geographic position (measured in terms of geodesic path length on the WGS-84 ellipsoid) and the RMSE in time of the declared LPE assigned to that launch point/time.

Calculation: From the calculation of launch point completeness, use the time-dependent set of real ballistic missile launches that have occurred by time t in the scenario, and use the result of the gated unique assignment of declared LPEs to true launches.

The TAMD CRD specifies that LPE accuracy measures are to be scored at certain event-driven times $t_{ScoreLPEAcc_{i,n}}$, which may be specific to the launch event i and run n being assessed. Also, it is recommended that scoring results be presented in terms of percentile statistics, rather than as averages. The recommended roll-up procedure for these attribute measures is thus to construct a cumulative distribution of all values of the instantaneous metric scored for each participant m, true launch i, and run n, at the scoring time $t_{ScoreLPEAcc_{i,n}}$.

The true launch event-specific LPE accuracy measures for true launch i, and participant m, in run n, are the LPE Position Accuracy $PosEr_{i,m,n}$ and LPE Time Accuracy $TimeEr_{i,m,n}$, defined for any (i,m,n) for which participant m holds an LPE assigned to launch i in run n at time $t_{ScoreLPEAcc_{i,n}}$, as follows:

$$\text{LPEPosEr}_{i,m,n} = V(\mathbf{E}_{i,m,n}, \mathbf{E}_{truth,i,n})$$
$$\text{LPETimeEr}_{i,m,n} = \text{LTime}_{i,m,n} - \text{LTime}_{truth,i,n}$$

where V() is the geodesic path length between two positions on the WGS-84 ellipsoid (computed, for example, by the Vincenty iterative method), $\mathbf{E}_{i,n,m}$ is the vector of latitude/longitude position coordinates of the LPE held by participant m and assigned to launch i at time $t_{ScoreLPEAcc_{i,n}}$ during run n, and $\mathbf{E}_{truth,i,n}$ is the vector of true launch point latitude/longitude position coordinates for launch i in that run; and where launch time estimate of the same LPE is $\text{LTime}_{i,m,n}$ and the time of the true launch is $\text{LTime}_{truth,i,n}$.

Histogram the statistical densities and plot the cumulative distributions of the quantities $\text{LPEPosEr}_{i,m,n}$ and $\text{LPETimeEr}_{i,m,n}$, for the entire set of evaluations of these metrics. Extract percentile statistics of interest (and averages, if required) from the cumulative distributions. If averages of LPE Position and Time Accuracy over user-specified intervals of time are also required, an averaging procedure analogous to that detailed above for Track Accuracy is recommended.

Accuracy: LPE Position and Time Uncertainty Consistency

Definition: As assessed for a particular real ballistic missile launch, the mean Chi-squared statistic of the LPE time estimate and the proportions of occurrences in which the true position error of the declared LPE is within the 95th percentile uncertainty areas reported with the LPE.

Remark: Calculation of the mean Chi-squared statistic requires variance information on launch time estimates. If such data are unavailable, then the time uncertainty consistency metric is not defined.

Calculation: Initialize $m_{95}(t_{ScoreLPE}) = 0$ and $\chi^2_{Time\_Aver\_0}(t_{ScoreLPE}) = 0$.

For each participant m and true launch i, at each scheduled LPE scoring time $t_{ScoreLPE}$ of each run, use the $n(t_{ScoreLPE})$ and $TE_n = \text{TimeEr}_{i,m,n}(t_{ScoreLPE})$ results from calculation of LPE position and time accuracy and use the launch time estimate t_i with variances $\sigma_{t\_i}^2$ and launch point estimate latitude and longitude coordinates (ϕ_i, λ_i) with 95th percentile ellipse major and minor axes, a_{maj_i} and a_{min_i}, and azimuth orientation of the major axis β_i to perform the following recursive calculations:

$$m_{95}(t_{ScoreLPE}) = \begin{cases} m_{95}(t_{ScoreLPE}) + 1 & \text{if } \left[\dfrac{R_{i,j}\cos(\alpha_{i,j} - \beta_i)}{a_{maj\_i}}\right]^2 + \left[\dfrac{R_{i,j}\sin(\alpha_{i,j} - \beta_i)}{a_{min\_i}}\right]^2 < 1 \\ m_{95}(t_{ScoreLPE}) & \text{otherwise} \end{cases}$$

$$\chi^2_{Time\_Aver\_n}(t_{ScoreLPE}) = \chi^2_{Time\_Aver\_n-1}(t_{ScoreLPE}) + \frac{1}{n}\left[\frac{TE_n^{\ 2}}{Var_{LPETime\_n}} - \chi^2_{Time\_Aver\_n-1}(t_{ScoreLPE})\right]$$

where $Var_{LPETime\_n} = \sigma_{t\_i}^2$ is the claimed time estimate variance for the assigned LPE at instance n, where $m_{95}(t_{ScoreLPE})$ is the count of occurrences when the position estimate error is within the 95th percentile uncertainty areas, where $\chi^2_{Time\_Aver\_n}(t_{ScoreLPE})$ is the mean Chi-squared statistic of the time estimate, and where $R_{i,j}$ and $\alpha_{i,j}$ are respectively the geodesic path length and azimuth angle from LPE to true launch point as used in the assignment algorithm of Section 3.3.4.2.

After completing the above recursion for participant m and object i of interest, plot $n(t_{ScoreLPE})$ as a function of $t_{ScoreLPE}$ (this represents the number of runs for which participant m holds an LPE assigned to launch i at the given scoring time). For time segments where $n(t_{ScoreTrk}) > 0$, set $PosCon_{LPE,i,m}(t_{ScoreLPE}) = [m_{95}(t_{ScoreLPE}) / n(t_{ScoreLPE})]$, and $\chi^2_{LPETime,i,m}(t_{ScoreLPE}) = \chi^2_{Time\_Aver\_n}(t_{ScoreLPE})$.

Plot $PosCon_{LPE,i,m}(t)$ and $\chi^2_{LPETime,i,m}(t)$ as functions of time t into the scenario for each participant m, each launch i of interest, and each scoring time interval within which the statistic has a defined value.

Higher-level roll-ups of this MOP may be computed as required by the user.

Accuracy: IPP Position and Time Accuracy

Definition: As assessed for a reportable object thread true impact, the RMSE in geographic position (in terms of geodesic path length on the WGS-84 ellipsoid) and the RMSE in time of the declared IPP assigned to that impact point/time.

Calculation: From the calculation of object track completeness, use the time-dependent set of real objects meeting reporting criteria. Cross-reference their object numbers to ground impact points of truth trajectories having (successive) primary child inheritance (i.e, impacts of currently reportable object threads).

The TAMD CRD specifies that IPP accuracy measures are to be scored at certain event-driven times $t_{ScoreIPPAcc_{j,n}}$, which may be specific to the reportable object thread j and run n being assessed. Also, it is recommended that scoring results be presented in terms of percentile statistics, rather than as averages. The recommended roll-up procedure for these attribute measures is thus to construct a cumulative distribution of all values of the instantaneous metric

scored for each participant m, impacting reportable object j, and run n, at the scoring time $t_{ScoreLPEAcc_{j,n}}$.

The reportable object thread-specific IPP accuracy measures for reportable object thread j, and participant m, in run n, are the IPP Position Accuracy $IPPPosEr_{j,m,n}$ and IPP Time Accuracy $IPPTimeEr_{j,m,n}$, defined for any (j,m,n) for which participant m holds an IPP assigned to the true impact of currently reportable object thread j in run n at time $t_{ScoreIPPAcc_{j,n}}$, as follows:

$$IPPPosEr_{j,m,n} = V(\mathbf{D}_{j,m,n}, \mathbf{D}_{truth,j,n})$$
$$IPPTimeEr_{j,m,n} = ITime_{j,m,n} - ITime_{truth,j,n}$$

where V() is the geodesic path length (computed, for example, by the Vincenty iterative method), $\mathbf{D}_{j,n,m}$ is the vector of latitude/longitude position coordinates of the IPP held by participant m and assigned to impacting currently reportable object thread j at time $t_{ScoreIPPAcc_{j,n}}$ during run n, and $\mathbf{D}_{truth,j,n}$ is the vector of true impact point latitude/longitude position coordinates for reportable object thread j in that run; and where impact time estimate of the same IPP is $ITime_{j,m,n}$ and the time of the true imapct is $ITime_{truth,j,n}$.

Histogram the statistical densities and plot the cumulative distributions of the quantities $IPPPosEr_{j,m,n}$ and $IPPTimeEr_{j,m,n}$, for the entire set of evaluations of these metrics. Extract percentile statistics of interest (and averages, if required) from the cumulative distributions. If averages of IPP Position and Time Accuracy over user-specified intervals of time are also required, an averaging procedure analogous to that detailed above for Track Accuracy is recommended.

Accuracy: IPP Position and Time Uncertainty Consistency

Definition: As assessed for a particular real object meeting reporting criteria, the mean Chi-squared statistic of the IPP time estimate and the proportions of occurrences in which the true position error of the declared IPP is within the 95th percentile uncertainty area reported with the IPP.

Remark: Calculation of the mean Chi-squared statistic requires variance information on impact time estimates. If such data are unavailable, then the time uncertainty consistency metric is not defined.

Calculation: Initialize $m_{95}(t_{ScoreIPP}) = 0$ and $\chi^2_{Time\_Aver\_0}(t_{ScoreIPP}) = 0$.

For each participant m and each impact i (of a currently reportable object thread), at each scheduled IPP scoring time $t_{ScoreIPP}$ of each run, use the $n(t_{ScoreIPP})$ and $TE_n = TimeEr_{i,m,n}(t_{ScoreIPP})$ results from calculation of IPP position and time accuracy and use the impact time prediction t_i with variances $\sigma_{t\_i}{}^2$ and

impact point prediction latitude and longitude coordinates (ϕ_i, λ_i) with 95th percentile ellipse major and minor axes, a_{maj_i} and a_{min_i}, and azimuth orientation of the major axis β_i to perform the following recursive calculations:

$$m_{95}(t_{ScoreIPP}) = \begin{cases} m_{95}(t_{ScoreIPP}) + 1 & \text{if } \left[\dfrac{R_{i,j}\cos(\alpha_{i,j} - \beta_i)}{a_{maj\_i}}\right]^2 + \left[\dfrac{R_{i,j}\sin(\alpha_{i,j} - \beta_i)}{a_{min\_i}}\right]^2 < 1 \\ m_{95}(t_{ScoreIPP}) & \text{otherwise} \end{cases}$$

$$\chi^2_{Time\_Aver\_n}(t_{ScoreIPP}) = \chi^2_{Time\_Aver\_n-1}(t_{ScoreIPP}) + \frac{1}{n}\left[\frac{TE_n^{\ 2}}{Var_{IPPTime\_n}} - \chi^2_{Time\_Aver\_n-1}(t_{ScoreIPP})\right],$$

where $Var_{IPPTime\_n} = \sigma_{t\_i}^2$ is the claimed impact time prediction variance for the assigned IPP at instance n, where $m_{95}(t_{ScoreIPP})$ is the count of occurrences when the position estimate error is within the 95th percentile uncertainty areas, where $\chi^2_{Time\_Aver\_n}(t_{ScoreIPP})$ is the mean Chi-squared statistic of the time estimate, and where $R_{i,j}$ and $\alpha_{i,j}$ are respectively the geodesic path length and azimuth angle from IPP to true impact point as used in the IPP gating procedure of Section 3.3.4.3.

After completing the above recursion for participant m and object i of interest, plot $n(t_{ScoreIPP})$ as a function of $t_{ScoreIPP}$ (this represents the number of runs for which participant m holds an IPP assigned to impact i at the given scoring time). For time segments where $n(t_{ScoreIPP}) > 0$, set $PosCon_{IPP,i,m}(t_{ScoreIPP})$ = $[m_{95}(t_{ScoreIPP}) / n(t_{ScoreIPP})]$, and $\chi^2_{IPPTime,i,m}(t_{ScoreIPP}) = \chi^2_{Time\_Aver\_n}(t_{ScoreIPP})$.

Plot $PosCon_{LPE,i,m}(t)$ and $\chi^2_{LPETime,i,m}(t)$ as functions of time t into the scenario for each participant m, each impact i of interest, and each scoring time interval within which the statistic has a defined value.

Higher-level roll-ups of this MOP may be computed as required by the user.

Timeliness: Track Initiation Delay

Definition: Delay from birth of a distinct reportable ballistic missile object (i.e., launch time for a unitary ballistic missile or a full launch stack, otherwise time of separation, release, breakup, segmentation, etc., that gives rise to the distinct reportable object), until the first time at which the object has a valid declared track.

Calculation: Determine the time-dependent set of real objects that meet reporting criteria, and the time of birth of those objects.

Whenever the participating unit declares a new firm track or receives a new network track, this should trigger an unscheduled scoring for track initiation time.

For each participant m and each run n, when there is a new track and at each scheduled track scoring time $t_{ScoreTrk}$, identify the gated unique assignment of declared tracks (including the new track) to the set of real object representations, according to the invoked metrics assignment method. On each run, record the first scoring time, $t_{i,m,n\_FirstTrack}$, that a declared track is assigned to object i meeting reporting criteria. If no declared track is assigned to object i prior to the scenario end time, then set $t_{i,m,n\_FirstTrack}$ equal to the scenario end time.

Histogram the statistical density and plot the cumulative distribution of $TID_{i,m,n} = (t_{i,m,n\_FirstTrack} - t_{i,n\_Birth})$ values across all evaluations of $t_{i,m,n\_FirstTrack}$, where t_{i,n_Birth} is the true time of birth of that object. Extract percentile statistics of interest (and averages, if required) from the cumulative distributions. Alternatively, the following direct formula for the average Track Initiation Delay over runs, reportable objects, and participants, may be applied:

$$TID = \frac{1}{M}\sum_{m=1}^{M}\frac{\sum_{i=1}^{L}\sum_{n=1}^{N} NRT_i \cdot (t_{i,m,nFirstTrack} - t_{i,nBirth})}{\sum_{i=1}^{L} NRT_i},$$

where NRT_i is the number of runs in which reportable object i appears (allowing for the possibility that the launch of some objects may be prevented by destruction of launch sites during certain runs).

Timeliness: LPE Delay

Definition: Delay from launch of a ballistic missile or LCR until the first launch point estimate is assigned to that true launch.

Calculation: Whenever the participating unit declares or receives a new LPE, this should trigger an unscheduled scoring for LPE delay.

For each participant m and each run n, when there is a new LPE and at each scheduled LPE scoring time $t_{ScoreLPE}$, identify a gated unique assignment of declared launch point/time estimates to the set of real launches of ballistic missiles that have occurred by that time into the scenario (following the assignment procedure described for computing LPE completeness). On each run, record the first scoring time, $t_{i,m,n\_FirstLPE}$, that a declared LPE is assigned to true launch i. If no declared LPE is assigned to true launch i prior to the scenario end time, then set $t_{i,m,n\_FirstLPE}$ equal to the scenario end time.

Histogram the statistical density and plot the cumulative distribution of $LD_{i,m,n} = (t_{i,m,n\_FirstLPE} - t_{i,n\_Launch})$ values across all evaluations of $t_{i,m,n\_FirstLPE}$, where t_{i,n_Launch} is the true launch time of launch i. Extract percentile statistics of interest (and averages, if required) from the cumulative distributions. Alternatively, the following direct formula for the average LPE Delay over runs, launches, and participants, may be applied:

$$LD = \frac{1}{M}\sum_{m=1}^{M}\frac{\sum_{i=1}^{LL}\sum_{n=1}^{N} NRL_i \cdot (t_{i,m,nFirstLPE} - t_{i,nLaunch})}{\sum_{i=1}^{LL} NRL_i},$$

where NRL_i is the number of runs in which launch i occurs.

Timeliness: Booster Burnout Estimate Delay

Definition: Delay from burnout of the final booster stage of a ballistic missile until the first valid booster burnout estimate with state vector is provided.

Calculation: First, determine the set of real ballistic missile final-stage booster burnout events that have occurred by time t in the scenario and the time/position/velocity states for those final burnout events, as well as the true launch events for the corresponding object threads.

Whenever the participating unit declares or receives a new booster burnout estimate, this should trigger an unscheduled scoring for booster burnout estimate delay.

For each participant m and each run n, when there is a new booster burnout estimate and at scheduled track scoring times $t_{ScoreTrk}$, identify the gated unique assignment of the entire set of initial and refined booster burnout time/position/velocity states, with error covariance, held by the unit to the set of real ballistic missile final-stage booster burnout events that have occurred by that time in the scenario. On each run, record the first scoring time, $t_{i,m,n\_FirstBstrBO}$, that a declared booster burnout event is assigned to the true final-stage booster burnout of the ballistic missile from launch point i. (Follow the procedure delineated below.) If no declared booster burnout estimate is assigned to the true event for the ballistic missile from launch point i prior to the scenario end time, then set $t_{i,m,n\_FirstBstrBO}$ equal to the scenario end time.

Histogram the statistical density and plot the cumulative distribution of $BBD_{i,m,n} = (t_{i,m,n\_FirstBstrBO} - t_{i,n\_BstrBO})$ values across all evaluations of $t_{i,m,n\_FirstBstrBO}$, where t_{i,n_BstrBO} is the true final stage booster burnout time associated with the ballistic missile from launch i. Extract percentile statistics of interest (and averages, if required) from the cumulative distributions. Alternatively, the

following direct formula for the average Booster Burnout Estimate Delay over runs, true final stage burnout events, and participants, may be applied:

$$BBD = \frac{1}{M}\sum_{m=1}^{M}\frac{\sum_{i=1}^{LB}\sum_{n=1}^{N} NRB_i \cdot (t_{i,m,nFirstBstrBO} - t_{i,nBstrBO})}{\sum_{i=1}^{LB} NRB_i},$$

where NRB_i is the number of runs in which the event of final stage booster burnout from launch i occurs, and LB is the number of true final ballistic missile final stage burnout events over the duration of the scenario.

Booster Burnout Estimate Gating and Assignment: Booster burnout estimates are gated and assigned directly to true ballistic missile final booster stage burnout events. The gating and assignment costs are to be computed as follows:

When there is a new booster burnout estimate and at each scheduled track scoring time of each run, obtain from the participating unit the set of initial and refined booster burnout estimate time/position/velocity states, with 7-state error covariance. (If no uncertainty information is provided/available for the booster burnout time estimate, use $\sigma_t = G_t/3$ for time gate value G_t and set the time/position and time/velocity cross covariance terms to zero.)

First, gate the booster burnout estimate i with respect to true time of final booster stage burnout event j by $|t_i - t_j| < G_t$ (e.g., $G_t = 30.0$ sec).

If the time gating above is passed, compute the 7-degree-of-freedom chi-squared distance $C(i,j) = \chi^2_{i,j} = (\mathbf{x}_i - \mathbf{x}_j)^T \mathbf{P}_i^{-1} (\mathbf{x}_i - \mathbf{x}_j)$ between the time/position/velocity states $\mathbf{x}_i$ and corresponding 7x7 error covariance $\mathbf{P}_i$ of the booster burnout estimate and the true time/position/velocity states $\mathbf{x}_j$ of final booster stage burnout event j, and gate this value by $\chi^2_{i,j} < G$ (e.g., G = 35.3, corresponding to a 0.99999 probability gate for seven degrees of freedom).

Across all the booster burnout estimates/true final stage booster burnout combinations (i,j) that are not gated out by any of the above criteria, perform an optimal unique assignment with the C(i,j) values from the above formula, setting the guard value equal to G (cf. SIAP SE TF Technical Report 2001-003, Section 3.3).

Timeliness: IPP Warning Time

Definition: For a particular reportable object thread impact, the true time of impact minus the first time a ground impact point prediction is assigned to that true impact.

Calculation: Whenever the participating unit declares or receives a new IPP, this should trigger an unscheduled scoring for IPP time.

For each participant m and each run n, when there is a new IPP and at each scheduled IPP scoring time $t_{ScoreIPP}$, identify the gated unique assignment of declared IPPs to the set of true currently reportable object thread impacts. On each run, record the first scoring time, $t_{j,m,n\_FirstIPP}$, that true currently reportable object thread impact j has a valid declared IPP. If true reportable object thread impact j has had no valid IPP prior to the true impact time, then set $t_{j,m,n\_FirstIPP}$ equal to the true impact time $t_{j,n\ True}$.

Histogram the statistical density and plot the cumulative distribution of $IWT_{j,m,n} = (t_{j,m,n\_FirstIPP} - t_{j,n\_True})$ values across all evaluations of $t_{j,m,n\_FirstIPP}$. Extract percentile statistics of interest (and averages, if required) from the cumulative distributions. Alternatively, the following direct formula for the average IPP Warning Time over runs, true impacts, and participants, may be applied:

$$IWT = \frac{1}{M}\sum_{m=1}^{M}\frac{\sum_{j=1}^{LI}\sum_{n=1}^{N} NRI_j \cdot (t_{j,m,nFirstIPP} - t_{j,nTrue})}{\sum_{j=1}^{LI} NRI_j},$$

where NRI_j is the number of runs in which the impact of reportable object thread j occurs, and LI is the total number of reportable object thread impacts which can occur in the scenario.

Correctness: Booster Typing Correctness

Definition: As assessed for a particular real ballistic missile launch, the match between the booster type probability (or single type declaration) vector reported with a valid LPE and the true booster type for that launch.

Remark: This definition is framed in terms of probability vectors in order to allow assessment of tracking systems which process typing data probabilistically. For systems that presently have this capability, the metric is to be evaluated as written. However, some legacy systems may not process typing probabilities, and Link 16 presently does not support reporting of probability fields through J3.0 LPE or J3.6 Space Track messages. Therefore, it is intended for these legacy

systems, and for remote data assessed through J3.0 LPE messages cross referenced to J3.6 track numbers, that the probability vectors entering into this definition represent only hard classification information (the component of the vector is 1 for the type being represented, 0 for all other types). If, on the other hand, no typing information is provided at all (i.e., a "no statement" response), then the typing probability vector should be set to a zero vector in the evaluation of this metric.

Calculation: From the calculation of launch point completeness, use the time-dependent set of real ballistic missile launches that have occurred by time t in the scenario, and use the result of the gated assignment of declared LPE to true launches.

The TAMD CRD specifies that Booster Typing Correctness is to be scored at certain event-driven times $t_{ScoreBstMatch_{i,m,n}}$, which may be specific to the launch event i participant m and run n being assessed. Also, it is recommended that scoring results be presented in terms of percentile statistics, rather than as averages. The recommended roll-up procedure for these attribute measures is thus to construct a cumulative distribution of all values of the instantaneous metric scored for each participant m, true launch i, and run n, at the scoring time $t_{ScoreBstMatch_{i,m,n}}$. However, a complete procedure for averaging the instantaneous metric, as evaluated at the usual LPE scoring times $t_{ScoreLPE_k}$, over time (and participants, launches, and runs), will also be provided in case the user requires it.

The true launch event-specific Booster Typing Correctness measure $BstMatch_{i,m,n}$ for true launch i, and participant m, in run n is defined for any (i,m,n) for which participant m holds an LPE assigned to launch i, as follows:

$$BstMatch_{i,m,n} = \mathbf{BT}_n(t_{ScoreBstMatch_{i,m,n}}) \bullet \mathbf{BT}_{truth_{i,n}}$$

where $\mathbf{BT}_n(t_{ScoreBstMatch_{i,m,n}})$ is the booster type probability vector associated with (i.e., reported in or cross referenced from) the assigned LPE in run n at time $t_{ScoreBstMatch_{i,m,n}}$, where $\mathbf{BT}_{truth_{i,n}}$ is the true booster type vector for that launch i (i.e., $\mathbf{BT}_{truth_{i,n}}$ has a value of 1.0 in the element for the true booster type and values of 0.0 for all other elements), and where $\mathbf{BT}_n(t_{ScoreBstMatch_{i,m,n}}) \bullet \mathbf{BT}_{truth_{i,n}}$ is their dot product.

Histogram the statistical densities and plot the cumulative distributions of the quantities $BstMatch_{i,m,n}$, for the entire set of evaluations of the metric. Extract percentile statistics of interest (and averages, if required) from the cumulative distributions. These results should be used to assess compliance with TAMD CRD EW requirements.

An alternative approach to rolling up the instantaneous metric

$$\mathrm{BstMatch}_{i,m,n}(t_{\mathrm{ScoreLPE}_k}) = \mathbf{BT}_n(t_{\mathrm{ScoreLPE}_k}) \bullet \mathbf{BT}_{\mathrm{truth}_{i,n}},$$

where t_{ScoreLPE_k} is the usual k^{th} scheduled LPE scoring time, over some user-specified interval(s) of interest, will now be described.

The launch-specific Booster Typing Correctness for launch i, participant m, and scoring time t_{ScoreTrk_k} are defined as appropriate averages over runs for which the scoring time is within an interval of interest:

$$\mathrm{BstMatch}_{i,m}(t_{\mathrm{ScoreLPE}_k}) = \frac{\sum_{n=1}^{N} \mathrm{WBTC}_{i,m,n,k} \mathrm{BstMatch}_{i,m,n}(t_{\mathrm{ScoreLPE}_k})}{\sum_{n=1}^{N} \mathrm{WBTC}_{i,m,n,k}},$$

where the Boolean variable $\mathrm{WBTC}_{i,m,n,k}$ is set to 1 if t_{ScoreLPE_k} is within a time interval of interest for launch in run n and participant m holds an LPE assigned to launch i at t_{ScoreLPE_k}, and set to 0 otherwise.

The run averages may also be computed by the following recursive procedure, performed at each scoring time for a particular launch and participant. (Subscripts identifying launch, participant, and scoring time increment are suppressed in the recursion formulae, since the key variables are re-initialized at the beginning of each recursion.)

For each participant m, at each scheduled LPE scoring time t_{ScoreLPE} of each run, if a LPE is assigned to true launch event i at t_{ScoreLPE}, increment $n=n(t_{\mathrm{ScoreLPE}})$ (i.e., n gets n+1) for that scheduled LPE scoring time and true launch, and perform the following recursion update:

$$\begin{aligned}\mathrm{BstMatch}_{\mathrm{Aver\_n}}(t_{\mathrm{ScoreLPE}}) &= \mathrm{BstMatch}_{\mathrm{Aver\_n-1}}(t_{\mathrm{ScoreLPE}}) \\ &\quad + \frac{1}{n}\left[\mathbf{BT}_n(t_{\mathrm{ScoreLPE}}) \bullet \mathbf{BT}_{\mathrm{truth}} - \mathrm{BstMatch}_{\mathrm{Aver\_n-1}}(t_{\mathrm{ScoreLPE}})\right]\end{aligned}$$

where $\mathbf{BT}_n(t_{\mathrm{ScoreLPE}})$ is the booster type probability vector associated with (i.e., reported in or cross referenced from) the assigned LPE at instance n, where $\mathbf{BT}_{\mathrm{truth}}$ is the true booster type vector for that launch point i (i.e., $\mathbf{BT}_{\mathrm{truth}}$ has a value of 1.0 in the element for the true booster type and values of 0.0 for all other elements), where $\mathbf{BT}_n(t_{\mathrm{ScoreLPE}}) \bullet \mathbf{BT}_{\mathrm{truth}}$ is their dot product, and where

$BstMatch_{Aver\_0}(t_{ScoreLPE}) = 0$. If the conditions for incrementing the index $n(t_{ScoreLPE})$ are not met, then go to the next run without updating any variables.

After completing the above recursion for participant m and true launch i of interest, set $NRL_{i,m}(t_{ScoreLPE_k}) = n(t_{ScoreLPE})$ (this represents the number of runs included in the average, and is also equal to $\sum_{n=1}^{N} WBTC_{i,m,n,k}$).

The instantaneous average Booster Typing Correctness $BstMatch_m(t_{ScoreLPE_k})$ for participant m at k[th] LPE scoring time is defined as the following weighted average over true launches:

$$BstMatch_m(t_{ScoreLPE_k}) = \frac{\sum_{i=1}^{LL} NRL_{i,m}(t_{ScoreLPE_k})BstMatch_{i,m}(t_{ScoreLPE_k})}{\sum_{i=1}^{LL} NRL_{i,m}(t_{ScoreLPE_k})}.$$

The instantaneous average Booster Typing Correctness over launches and participants is then defined as the weighted average of $BstMatch_m(t_{ScoreLPE_k})$ over all participants for which it is defined:

$$BstMatch(t_{ScoreLPE_k}) = \frac{\sum_{m=1}^{M}\sum_{i=1}^{LL} NRL_{i,m}(t_{ScoreLPE_k})BstMatch_{i,m}(t_{ScoreLPE_k})}{\sum_{m=1}^{M}\sum_{i=1}^{LL} NRL_{i,m}(t_{ScoreLPE_k})},$$

and finally, the average Booster Track Completeness over launches, participants, and time is given by:

$$BstMatch = \frac{\sum_{k=1}^{K}\sum_{m=1}^{M}\sum_{i=1}^{LL} NRL_{i,m}(t_{ScoreLPE_k})BstMatch(t_{ScoreLPE_k})\Delta t_{ScoreLPE_k}}{\sum_{k=1}^{K}\sum_{m=1}^{M}\sum_{i=1}^{LL} NRL_{i,m}(t_{ScoreLPE_k})\Delta t_{ScoreLPE_k}}.$$

Correctness: Post-Boost Object Classification Correctness

Definition: As assessed for a particular real post-boost object meeting reporting criteria, the match between the reported object classification probability vector (or single classification declaration) associated with a valid track and the true classification of the post-boost object.

Remark: This definition is framed in terms of probability vectors in order to allow assessment of tracking systems which process PBC data probabilistically. For systems that presently have this capability, the metric is to be evaluated as written. However, some legacy systems may not process PBC probabilities, and Link 16 presently does not support reporting of probability fields through J3.6 Space Track messages. Therefore, it is intended for these legacy systems and for remote data assessed through J3.6 messages that the probability vectors entering into this definition represent only hard classification information (the component of the vector is 1 for the object class being represented, 0 for all other classes). If, on the other hand, no PBC information is provided at all (i.e., a "no statement" response), then the classification probability vector should be set to a zero vector in the evaluation of this metric.

Calculation: Determine the set of real objects meeting track and post-boost object classification reporting criteria.

The TAMD CRD specifies that Post-Boost Object Classification Correctness is to be scored at certain event-driven times $t_{ScoreObjMatch_{i,m,n}}$, which may be specific to the reportable post-boost object i participant m and run n being assessed. Also, it is recommended that scoring results be presented in terms of percentile statistics, rather than as averages. The recommended roll-up procedure for these attribute measures is thus to construct a cumulative distribution of all values of the instantaneous metric scored for each participant m, true launch i, and run n, at the scoring time $t_{ScoreObjMatch_{i,m,n}}$.

The reportable post-boost-object-specific classification measure $ObjMatch_{i,m,n}$ for reportable post-boost object i, and participant m in run n, is defined for any (i,m,n) for which participant m holds a track assigned to post-boost object i, as follows:

$$\mathrm{ObjMatch}_{i,m,n} = \mathbf{OT}_n(t_{ScoreBstMatch_{i,m,n}}) \bullet \mathbf{OT}_{truth_{i,n}}$$

where $\mathbf{OT}_n(t_{ScoreObjMatch_{i,m,n}})$ is the post-booster object classification probability vector reported for the assigned track in run n at time $t_{ScoreObjMatch_{i,m,n}}$, where $\mathbf{OT}_{truth_{i,n}}$ is the true booster type vector for that launch i (i.e., $\mathbf{OT}_{truth_{i,n}}$ has a value of 1.0 in the element for the true object class and values of 0.0 for all other elements), and where $\mathbf{OT}_n(t_{ScoreObjMatch_{i,m,n}}) \bullet \mathbf{OT}_{truth_{i,n}}$ is their dot product.

Histogram the statistical densities and plot the cumulative distributions of the quantities $ObjMatch_{i,m,n}$, for the entire set of evaluations of the metric. Extract percentile statistics of interest (and averages, if required) from the cumulative distributions. If averages of Post-Boost Object Classification Correctness over

user-specified intervals of time are also required, an averaging procedure analogous to that detailed above for Booster Typing Correctness is recommended.

Correctness: IPP Object Classification Correctness

Definition: As assessed for a particular reportable object thread impact, the match between the impacting object classification probability vector (or single classification declaration) associated with the declared IPP assigned to that impact point/time and the true classification of the impacting object.

Remark: This definition is framed in terms of probability vectors in order to allow assessment of tracking systems which process impacting object classification data probabilistically. For systems that presently have this capability, the metric is to be evaluated as written. However, some legacy systems may not process classification probabilities, and Link 16 presently does not support reporting of probability fields through J3.0 IPP or J3.6 Space Track messages. Therefore, it is intended for these legacy systems and for remote data assessed through J3.0 IPP messages cross referenced to J3.6 track numbers that the probability vectors entering into this definition represent only hard classification information (the component of the vector is 1 for the type being represented, 0 for all other types). If, on the other hand, no impacting object classification information is provided at all (i.e., a “no statement” response), then the classification probability vector should be set to a zero vector in the evaluation of this metric.

Calculation: From the calculation of impact point completeness, use the time-dependent set of true currently reportable object thread impacts, and use the result of the gated assignment of declared IPPs to true currently reportable object thread impacts.

The TAMD CRD is interpreted as specifying that IPP Object Classification Correctness be scored at certain event-driven times $t_{ScoreIPPClassMatch_{j,m,n}}$, which may be specific to the reportable object thread impact j participant m and run n being assessed. Also, it is recommended that scoring results be presented in terms of percentile statistics, rather than as averages. The recommended roll-up procedure for these attribute measures is thus to construct a cumulative distribution of all values of the instantaneous metric scored for each participant m, impacting reportable object thread j, and run n, at the scoring time $t_{ScoreIPPClassMatch_{j,m,n}}$.

The reportable post-boost-object-specific classification measure $IPPClassMatch_{j,m,n}$ for impacting reportable object thread j, and participant m in run n, is defined for any (j,m,n) for which participant m holds an IPP assigned to object thread j, as follows:

$$\text{IPPClassMatch}_{j,m,n} = \mathbf{IOT}_n(t_{\text{IPPClassMatch}_{j,m,n}}) \bullet \mathbf{OT}_{\text{truth}_{i,n}}$$

where $\mathbf{IOT}_n(t_{\text{IPPClassMatch}_{j,m,n}})$ is the impacting object classification probability vector associated with (i.e., reported in or cross referenced from) the IPP assigned to reportable object thread j in run n at time $t_{\text{ScoreObjMatch}_{i,m,n}}$, where $\mathbf{IOT}_{\text{truth}_{j,n}}$ is the true booster type vector for that launch i (i.e., $\mathbf{IOT}_{\text{truth}_{j,n}}$ has a value of 1.0 in the element for the true object class and values of 0.0 for all other elements), and where $\mathbf{IOT}_n(t_{\text{IPPClassMatch}_{ji,m,n}}) \bullet \mathbf{IOT}_{\text{truth}_{j,n}}$ is their dot product.

Histogram the statistical densities and plot the cumulative distributions of the quantities $\text{IPPClassMatch}_{j,m,n}$, for the entire set of evaluations of the metric. Extract percentile statistics of interest (and averages, if required) from the cumulative distributions. If averages of IPP Object Classification Correctness over user-specified intervals of time are also required, an averaging procedure analogous to that detailed above for Booster Typing Correctness is recommended.

Commonality: Track Commonality

Definition: Ratio of the number of valid declared tracks held by all participants, with consistent track numbers, kinematic data, and typing/classification data, to the total number of valid declared tracks.

Calculation:

Let $\text{NCT}_n(t)$ be the number of declared valid tracks held by all participants at time t during run n, such that

- each track is represented by the same track number, common to all participants
- time-aligned (by extrapolation or interpolation) position data on each track are the same for all participants, to within specified tolerances
- booster typing of a boost phase track or post-boost classification (as applicable) probability vectors are the same for all participants, to within specified tolerances.

Tolerances are intended to be flexible to accommodate different scenarios. It is recommended that a default position-only tolerance be set at 5 km separation in three-dimensional Euclidean distance. For booster typing and post-boost classification, "the same for all participants, to within specified tolerances" has the following interpretation. If $\mathbf{X}_m(t)$ and $\mathbf{X}_k(t)$ are the appropriate probability vectors ($\mathbf{X}$=$\mathbf{BT}$ for the boost phase, $\mathbf{OT}$ for the post-boost phase) associated by participants m and k with the same track number, then the probabilities are considered the same to within tolerance if the ratio

$$\frac{\| \mathbf{X}_m(t) - \mathbf{X}_k(t) \|^2}{\| \mathbf{X}_m(t) \|^2 + \| \mathbf{X}_k(t) \|^2}$$

is less than the specified tolerance value. The default tolerance in this case may be any positive number less than or equal to 1.0 (ensuring that "hard classification" probability vectors are scored as common only if they are identical). Default tolerances will be used unless the specific scenario or experiment should dictate otherwise. In connection with this and the following two commonality metrics, a "no statement" response for typing or classification will be interpreted as a uniform probability vector. That is, if participant m provides no typing/classification statement at time t, then $\mathbf{X}_m(t)$ is set equal to $(1/Q)\mathbf{1}$, where Q is the dimension of the vector space being used for typing/classification probabilities, and **1** is the vector in the same space with all unit entries. By this convention, a track with a "no statement" response may be regarded as common with a track indicating all typing/classification options equally likely.

Let $NS_n(t)$ be the number of valid declared tracks held by at least one participant at time t during run n (i.e., $NS_n(t)$ is the number of entries in the time-dependant set that is the union, across all participating units, of track numbers for declared tracks that are assigned to reportable objects). The Track Number Commonality at the kth track scoring time, $CMT(t_{ScoreTrk_k})$, is then computed by the following declared-track-weighted average over runs:

$$CMT(t_{ScoreTrk_k}) = \frac{\sum_{n=1}^{N} NCT_n(t_{ScoreTrk_k})}{\sum_{n=1}^{N} NS_n(t_{ScoreTrk_k})}$$

Track commonality may then be rolled up as an analogously weighted average over time, as follows:

$$CMT = \frac{\sum_{k=1}^{K}\sum_{n=1}^{N} NCT_n(t_{ScoreTrk_k})\Delta t_{ScoreLPE_k}}{\sum_{k=1}^{K}\sum_{n=1}^{N} NS_n(t_{ScoreTrk_k})\Delta t_{ScoreLPE_k}}.$$

Commonality: LPE Commonality

Definition: Ratio of the number of true launch points/times, with consistent booster typing/classification data, for which all participants hold a declared LPE assigned to the launch point, to the total number of valid declared LPEs.

Calculation: Let $\mathrm{NCLP}_n(t)$ be the number of true launch points for which at time t during run n every participant holds,

- a declared LPE assigned to that launch point/time (i.e., a valid LPE for that true launch event),
- every pair of these declared LPEs assigned to the same launch event has common position/time data to within specified tolerances, and
- booster type probability vectors are the same for all participants, to within specified tolerances.

Tolerances are intended to be flexible to accommodate different scenarios. It is recommended that a default position/time tolerance be set at 10 km surface distance (i.e., geodesic path length, computed for example by the Vincenty method) and 10 seconds time difference. For booster typing classification, "the same for all participants, to within specified tolerances" has the following interpretation. If $\mathbf{BT}_m(t_{ScoreLPE})$ and $\mathbf{BT}_k(t_{ScoreLPE})$ are the appropriate booster typing probability vectors associated by participants m and k for the same launch event, then the probabilities are considered the same to within tolerance if the ratio

$$\frac{\|\mathbf{BT}_m(t_{ScoreLPE})-\mathbf{BT}_k(t_{ScoreLPE})\|^2}{\|\mathbf{BT}_m(t_{ScoreLPE})\|^2+\|\mathbf{BT}_k(t_{ScoreLPE})\|^2}$$

is less than the specified tolerance value. The default tolerance in this case may be any positive number less than or equal to 1.0 (ensuring that "hard classification" probability vectors are scored as common only if they are identical). Default tolerances will be used unless the specific scenario or experiment should dictate otherwise. (See the calculation of Track Commonality for the convention for accounting for a "no statement" response.)

Let $\mathrm{NSLP}_n(t)$ be the number of valid declared LPEs held by at least one participant at time t in run n (i.e., $\mathrm{NSLP}_n(t)$ is the number of entries in the time-dependant set that is the union, across all participating units, of true launch events for which declared LPEs are assigned). The LPE Commonality at the LPE scoring time, $\mathrm{CMLP}(t_{ScoreLPE_k})$, is then computed by the following declared-LPE-weighted average over runs:

$$\mathrm{CMLP}(t_{ScoreLPE_k})=\frac{\sum_{n=1}^{N}\mathrm{NCLP}_n(t_{ScoreLPE_k})}{\sum_{n=1}^{N}\mathrm{NSLP}_n(t_{ScoreLPE_k})}$$

LPE commonality may then be rolled up as an analogously weighted average over time, as follows:

$$\mathrm{CMLP} = \frac{\sum_{k=1}^{K}\sum_{n=1}^{N} \mathrm{NCLP}_n(t_{\mathrm{ScoreLPE}_k})\Delta t_{\mathrm{ScoreLPE}_k}}{\sum_{k=1}^{K}\sum_{n=1}^{N} \mathrm{NSLP}_n(t_{\mathrm{ScoreLPE}_k})\Delta t_{\mathrm{ScoreLPE}_k}}.$$

Commonality: IPP Commonality

Definition: Ratio of the number of true impact points of reportable objects, with consistent impacting object classification data, for which all participants hold a declared IPP, to the total number of valid declared IPPs.

Calculation: Let $\mathrm{NCIP}_n(t)$ be the number of true currently reportable object thread impacts for which every participant at time t during run n holds,

- a declared IPP assigned to that impact point/time (i.e., a valid IPP for that true currently reportable object thread impact),
- every pair of these declared IPPs assigned to the same currently reportable object thread impact has common position/time data to within specified tolerances, and
- impacting object classification probability vectors are the same for all participants, to within specified tolerances.

Tolerances are intended to be flexible to accommodate different scenarios. It is recommended that a default position/time tolerance be set at 20 km surface distance (i.e., geodesic path length, computed for example by the Vincenty method) and 20 seconds time difference. For impacting object classification, "the same for all participants, to within specified tolerances" has the following interpretation. If $\mathbf{IOC}_m(t_{\mathrm{ScoreIPP}})$ and $\mathbf{IOC}_k(t_{\mathrm{ScoreIPP}})$ are the appropriate classsification probability vectors associated by participants m and k for the same currently reportable object thread impact, then the probabilities are considered the same to within tolerance if the ratio

$$\frac{\|\mathbf{IOC}_m(t_{\mathrm{ScoreIPP}}) - \mathbf{IOC}_k(t_{\mathrm{ScoreIPP}})\|^2}{\|\mathbf{IOC}_m(t_{\mathrm{ScoreIPP}})\|^2 + \|\mathbf{IOC}_k(t_{\mathrm{ScoreIPP}})\|^2}$$

is less than the specified tolerance value. The default tolerance in this case may be any positive number less than or equal to 1.0 (ensuring that "hard classification" probability vectors are scored as common only if they are identical). Default tolerances will be used unless the specific scenario or experiment should dictate otherwise. (See the calculation of Track Commonality for the convention for accounting for a "no statement" response.)

Let $\mathrm{NSIP}_n(t)$ be the number of valid declared IPPs held by at least one participant at time t in run n (i.e., $\mathrm{NSIP}_n(t)$ is the number of entries in the time-dependant set that is the union, across all participating units, of true currently reportable object thread impacts for which declared IPPs are assigned). The IPP

Commonality at the IPP scoring time, $\mathrm{CMIP}(t_{\mathrm{ScoreIPP}_k})$, is then computed as the following declared-IPP-weighted average over runs:

$$\mathrm{CMIP}(t_{\mathrm{ScoreIPP}_k}) = \frac{\sum_{n=1}^{N} \mathrm{NCIP}_n(t_{\mathrm{ScoreIPP}_k})}{\sum_{n=1}^{N} \mathrm{NSIP}_n(t_{\mathrm{ScoreIPP}_k})}$$

IPP commonality may then be rolled up as an analogously weighted average over time, as follows:

$$\mathrm{CMIP} = \frac{\sum_{k=1}^{K}\sum_{n=1}^{N} \mathrm{NCIP}_n(t_{\mathrm{ScoreIPP}_k})\Delta t_{\mathrm{ScoreIPP}_k}}{\sum_{k=1}^{K}\sum_{n=1}^{N} \mathrm{NSIP}_n(t_{\mathrm{ScoreIPP}_k})\Delta t_{\mathrm{ScoreIPP}_k}}.$$

Pairwise Cross-Platform Commonality History: Ratio of Non-Common Network Track Numbers

Definition: Number of active network track numbers that are different (additions or deletions) between pairs of participating units, divided by the number of network track numbers in the union of the two network track files.

Calculation: Let $\mathrm{TN}_{j,m}(t_{\mathrm{ScoreTrk}})$ and $\mathrm{TN}_{k,m}(t_{\mathrm{ScoreTrk}})$ be the sets of network track numbers held by participating units j and k (j not equal to k) at scheduled track scoring time t_{ScoreTrk} of run m. Let $f_N[A]$ be the function that returns the number of elements contained in set A. Then, the ratio of non-common track numbers between participating units j and k at t_{ScoreTrk} for run m is:

$$\mathrm{RNTN}_{j,k,m}(t_{\mathrm{ScoreTrk}}) = \frac{f_N\left[\mathrm{TN}_{j,m}(t_{\mathrm{ScoreTrk}}) \cup \mathrm{TN}_{k,m}(t_{\mathrm{ScoreTrk}}) - \mathrm{TN}_{j,m}(t_{\mathrm{ScoreTrk}}) \cap \mathrm{TN}_{k,m}(t_{\mathrm{ScoreTrk}})\right]}{f_N\left[\mathrm{TN}_{jm}(t_{\mathrm{ScoreTrk}}) \cup \mathrm{TN}_{k,m}(t_{\mathrm{ScoreTrk}})\right]}$$

For selected pairs of participating units of interest, average $\mathrm{RNTN}_{j,k,m}(t_{\mathrm{ScoreTrk}})$ over all the runs, i.e.,

$\mathrm{RNTN}_{j,k}(t_{\mathrm{ScoreTrk}}) = \frac{1}{N}\sum_{m=1}^{N} \mathrm{RNTN}_{j,k,m}(t_{\mathrm{ScoreTrk}})$, and plot $\mathrm{RNTN}_{j,k}(t)$ as a function of scenario time.

Also, for all participating unit pairs, all runs, and all scoring times, plot the statistical density and cumulative distribution of these $\mathrm{RNTN}_{j,k,m}(t_{\mathrm{ScoreTrk}})$ statistics.

Higher-level roll-ups of this MOP may be computed as required by the user.

Pairwise Cross-Platform Commonality History: Track State Estimate Differences

Definition: The Euclidean differences between position and velocity state estimates of tracks held by pairs of participating units, for tracks with the same active network track number.

Calculation: For each pair j,k of participating units, use the $TN_{j,m}(t_{ScoreTrk})$ and $TN_{k,m}(t_{ScoreTrk})$ sets determined for the rate of non-common track numbers. If relative gridlocking is invoked, then at each scoring time $t_{ScoreTrk}$ of run m, for tracks with common number n, translate the position/velocity state estimates of j and k in their local frames, $\mathbf{u}_{j,m,n}(t_{ScoreTrk})$ and $\mathbf{u}_{k,m,n}(t_{ScoreTrk})$, to WGS-84 earth centered, earth fixed Cartesian coordinates, $\mathbf{x}_{j,m,n}(t_{ScoreTrk})$ and $\mathbf{x}_{k,m,n}(t_{ScoreTrk})$, including compensation for the exact values of sensor latitude, longitude, altitude, and azimuth biases used in that series of runs, i.e.,

$$\mathbf{x}_{jm,n}(t_{Score}) = f\left[\mathbf{u}_{j,m,n}(t_{Score}); \Delta Lat_j, \Delta Long_j \Delta Alt_j \Delta Az_j\right]$$

$$\mathbf{x}_{k,m,n}(t_{Score}) = f\left[\mathbf{u}_{k,m,n}(t_{Score}); \Delta Lat_k, \Delta Long_k \Delta Alt_k \Delta Az_k\right]$$

where f is the six-state coordinate transformation from the local to the earth-fixed frame, and the Δs are the aforementioned biases, thus leaving only residual biases. (This step is not needed for geodetic absolute gridlock/registration).

For each scoring time, each run, each pair of participating units, and each network track number n in the set $TN_{j,m}(t_{ScoreTrk}) \cap TN_{k,m}(t_{ScoreTrk})$, record the following

$$\Delta P_{j,k,m,n}(t_{ScoreTrk}) = \sqrt{\left[P1_{j,m,n}(t_{ScoreTrk}) - P1_{k,m,n}(t_{ScoreTrk})\right]^2 + \left[P2_{j,m,n}(t_{ScoreTrk}) - P2_{k,m,n}(t_{ScoreTrk})\right]^2 + \left[P3_{j,m,n}(t_{ScoreTrk}) - P3_{k,m,n}(t_{ScoreTrk})\right]^2}$$

$$\Delta V_{j,k,m,n}(t_{ScoreTrk}) = \sqrt{\left[V1_{j,m,n}(t_{ScoreTrk}) - V1_{k,m,n}(t_{ScoreTrk})\right]^2 + \left[V2_{j,m,n}(t_{ScoreTrk}) - V2_{k,m,n}(t_{ScoreTrk})\right]^2 + \left[V3_{j,m,n}(t_{ScoreTrk}) - V3_{k,m,n}(t_{ScoreTrk})\right]^2}$$

where Pi and Vi are position and velocity components of $\mathbf{x}_{j,m,n}(t_{ScoreTrk})$ and $\mathbf{x}_{k,m,n}(t_{ScoreTrk})$.

At each scoring time $t_{ScoreTrk}$ for pairs of participating units of interest, compute and plot over the scenario time the root mean squared (RMS) position and velocity differences over all network track numbers n held in common and over all runs m, i.e.,

$$\Delta P_{RMS\_j,k}(t_{ScoreTrk}) = \sqrt{\frac{1}{N}\sum_{m=1}^{N}\left\{\frac{1}{f_N\left[TN_{j,m}(t_{ScoreTrk}) \cap TN_{k,m}(t_{ScoreTrk})\right]}\left[\sum_{n \in \{TN_{j,m}(t_{ScoreTrk}) \cap TN_{k,m}(t_{ScoreTrk})\}} \Delta P_{j,k,m,n}(t_{ScoreTrk})^2\right]\right\}}$$

$$\Delta V_{RMS\_j,k}(t_{ScoreTrk}) = \sqrt{\frac{1}{N}\sum_{m=1}^{N}\left\{\frac{1}{f_N\left[TN_{j,m}(t_{ScoreTrk}) \cap TN_{k,m}(t_{ScoreTrk})\right]}\left[\sum_{n \in \{TN_{j,m}(t_{ScoreTrk}) \cap TN_{k,m}(t_{ScoreTrk})\}} \Delta V_{j,k,m,n}(t_{ScoreTrk})^2\right]\right\}}$$

For all participating unit pairs, all network track numbers held in common by the pairs, all runs, and all scoring times, plot the statistical density and cumulative distribution of the $\Delta P_{j,k,m,n}(t_{ScoreTrk})$ and $\Delta V_{j,k,m,n}(t_{ScoreTrk})$ statistics. (Use random sampling if the data sets are too large to work with.)

Higher-level roll-ups of this MOP may be computed as required by the user.

Pairwise Cross-Platform Commonality History: Ratio of Non-Common Launch Points

Definition: Number of covered true ballistic missile launch points that are different (additions or deletions) between declared LPEs of pairs of participating units, divided by the number of true ballistic missile launch points covered in the union of the LPEs.

Calculation: From the calculation of launch point completeness, determine for each of a pair of participating units j and k on each run m the sets $LP_{j,m}(t_{ScoreLPE})$ and $LP_{k,m}(t_{ScoreLPE})$ of true ballistic missile launch points/times to which a declared LPE is assigned in the respective gated unique assignments. Let $f_N[A]$ be the function that returns the number of elements contained in set A. Then, the ratio of non-common launch points between participating units j and k at $t_{ScoreLPE}$ for run m is:

$$RNLP_{j,k,m}(t_{ScoreLPE}) = \frac{f_N\left[LP_{j,m}(t_{ScoreLPE}) \cup LP_{k,m}(t_{ScoreLPE}) - LP_{j,m}(t_{ScoreLPE}) \cap LP_{k,m}(t_{ScoreLPE})\right]}{f_N\left[LP_{j,m}(t_{ScoreLPE}) \cup LP_{k,m}(t_{ScoreLPE})\right]}$$

For selected pairs of participating units of interest, average $RNLP_{j,k,m}(t_{ScoreLPE})$ over all the runs, i.e.,

$RNLP_{j,k}(t_{ScoreLPE}) = \frac{1}{N}\sum_{m=1}^{N} RNLP_{j,k,m}(t_{ScoreLPE})$, and plot $RNLP_{j,k}(t)$ as a function of scenario time.

Also, for all participating unit pairs, all runs, and all scoring times, plot the statistical density and cumulative distribution of these $RNLP_{j,k,m}(t_{ScoreLPE})$ statistics.

Higher-level roll-ups of this MOP may be computed as required by the user.

Pairwise Cross-Platform Commonality History: LPE Differences

Definition: For pairs of participating units, the geographical position and time differences between LPEs held by the pair that are assign to the same true ballistic missile launch.

Calculation: For each pair j,k of participating units, use the $LP_{j,m}(t_{ScoreLPE})$ and $LP_{k,m}(t_{ScoreLPE})$ sets determined during the calculation of the ratio of non-common launch points.

For each schedule LPE scoring time $t_{ScoreLPE}$, each run m, each pair of participating units, and each true ballistic missile launch point n in the set $LP_{j,m}(t_{ScoreLPE}) \cap LP_{k,m}(t_{ScoreLPE})$, record the following

$$\Delta Pos_{j,k,m,n}(t_{ScoreLPE}) = V(\mathbf{E}_{j,m,n}, \mathbf{E}_{k,m,n})$$

$$\Delta Time_{j,k,m,n}(t_{ScoreLPE}) = abs\left[Time_{j,m,n}(t_{ScoreLPE}) - Time_{k,m,n}(t_{ScoreLPE})\right]$$

where V() is the geodesic path length (computed, for example, by the Vincenty method), $\mathbf{E}_{j,m,n} = \mathbf{E}_{j,m,n}(t_{ScoreLPE})$ and $\mathbf{E}_{k,m,n} = \mathbf{E}_{k,m,n}(t_{ScoreLPE})$ are the latitude/longitude position coordinate estimates from participants j and k respectively, and $Time_{j,m,n}(t_{ScoreLPE})$ and $Time_{k,m,n}(t_{ScoreLPE})$ are the time estimates of LPEs held respectively by participating units j and k assigned to true launch point n on run m.

At each scheduled scoring time $t_{ScoreLPE}$ for pairs of participating units of interest, compute and plot over the scenario time the root mean squared (RMS) position and time differences over all true launch points n covered in common and over all runs m, i.e.,

$$\Delta P_{RMS\_j,k}(t_{ScoreLPE}) = \sqrt{\frac{1}{N}\sum_{m=1}^{N}\left\{\frac{1}{f_N\left[LP_{j,m}(t_{ScoreLPE}) \cap LP_{k,m}(t_{ScoreLPE})\right]}\left[\sum_{n \in \{LP_{j,m}(t_{ScoreLPE}) \cap LP_{k,m}(t_{ScoreLPE})\}} \Delta Pos_{j,k,m,n}(t_{ScoreLPE})^2\right]\right\}}$$

$$\Delta Time_{RMS\_j,k}(t_{ScoreLPE}) = \sqrt{\frac{1}{N}\sum_{m=1}^{N}\left\{\frac{1}{f_N\left[LP_{j,m}(t_{ScoreLPE}) \cap LP_{k,m}(t_{ScoreLPE})\right]}\left[\sum_{n \in \{LP_{j,m}(t_{ScoreLPE}) \cap LP_{k,m}(t_{ScoreLPE})\}} \Delta Time_{j,k,m,n}(t_{ScoreLPE})^2\right]\right\}}$$

For all participating unit pairs, all true launch points covered in common by the pairs, all runs, and all scoring times, plot the statistical density and cumulative distribution of the $\Delta Pos_{j,k,m,n}(t_{ScoreLPE})$ and $\Delta Time_{j,k,m,n}(t_{ScoreLPE})$ statistics.

Higher-level roll-ups of this MOP may be computed as required by the user.

Pairwise Cross-Platform Commonality History: Ratio of Non-Common Impact Points

Definition: Number of covered true reportable object thread impacts that are different (additions or deletions) between declared IPPs of pairs of participating units, divided by the number of true reportable object thread impacts covered in the union of the IPPs.

Calculation: From the calculation of impact point completeness, determine for each of a pair of participating units j and k on each run m the sets $IP_{j,m}(t_{ScoreIPP})$ and $IP_{k,m}(t_{ScoreIPP})$ of true currently reportable object thread impact points/times to which a declared IPP is assigned in the respective gated unique assignments. Let $f_N[A]$ be the function that returns the number of elements contained in set A. Then, the ratio of non-common impact points between participating units j and k at $t_{ScoreLPE}$ for run m is:

$$RNIP_{j,k,m}(t_{ScoreIPP}) = \frac{f_N\left[IP_{j,m}(t_{ScoreIPP}) \cup IP_{k,m}(t_{ScoreIPP}) - IP_{j,m}(t_{ScoreIPP}) \cap IP_{k,m}(t_{ScoreIPP})\right]}{f_N\left[IP_{j,m}(t_{ScoreIPP}) \cup IP_{k,m}(t_{ScoreIPP})\right]}$$

For selected pairs of participating units of interest, average $RNIP_{j,k,m}(t_{ScoreIPP})$ over all the runs, i.e., $RNIP_{j,k}(t_{ScoreIPP}) = \frac{1}{N}\sum_{m=1}^{N} RNIP_{j,k,m}(t_{ScoreIPP})$, and plot $RNIP_{j,k}(t)$ as a function of scenario time.

Also, for all participating unit pairs, all runs, and all scoring times, plot the statistical density and cumulative distribution of these $RNIP_{j,k,m}(t_{ScoreIPP})$ statistics.

Higher-level roll-ups of this MOP may be computed as required by the user.

Pairwise Cross-Platform Commonality History: IPP Differences

Definition: For pairs of participating units, the geographical position and time differences between IPPs held by the pair that are assigned to the same true reportable object thread impact.

Calculation: For each pair j,k of participating units, use the $IP_{j,m}(t_{ScoreIPP})$ and $IP_{k,m}(t_{ScoreIPP})$ sets determined during the calculation of the ratio of non-common impact points.

For each schedule IPP scoring time $t_{ScoreIPP}$, each run m, each pair of participating units, and each true currently reportable object thread impact n in the set $IP_{j,m}(t_{ScoreIPP}) \cap IP_{k,m}(t_{ScoreIPP})$, record the following

$$\Delta Pos_{j,k,m,n}(t_{ScoreIPP}) = \frac{1}{2}\left[V(\mathbf{E}_{j,m,n},\mathbf{E}_{k,m,n}) + V(\mathbf{E}_{k,m,n},\mathbf{E}_{j,m,n})\right]$$

$$\Delta Time_{j,k,m,n}(t_{ScoreIPP}) = abs\left[Time_{j,m,n}(t_{ScoreIPP}) - Time_{k,m,n}(t_{ScoreIPP})\right]$$

where V() is the geodesic path length (computed, for example, by the Vincenty method), $\mathbf{E}_{j,m,n} = \mathbf{E}_{j,m,n}(t_{ScoreIPP})$ and $\mathbf{E}_{k,m,n} = \mathbf{E}_{k,m,n}(t_{ScoreIPP})$ are the latitude/longitude position coordinate estimates from participants j and k respectively, and $Time_{j,m,n}(t_{ScoreIPP})$ and $Time_{k,m,n}(t_{ScoreIPP})$ are the time estimates of IPPs held respectively by participating units j and k assigned to true currently reportable object thread impact n on run m.

At each scheduled scoring time $t_{ScoreIPP}$ for pairs of participating units of interest, compute and plot over the scenario time the root mean squared (RMS) position and time differences over all true currently reportable object thread impact points n covered in common and over all runs m, i.e.,

$$\Delta P_{RMS\_j,k}(t_{ScoreIPP}) = \sqrt{\frac{1}{N}\sum_{m=1}^{N}\left\{\frac{1}{f_N\left[IP_{j,m}(t_{ScoreIPP}) \cap IP_{k,m}(t_{ScoreIPP})\right]}\left[\sum_{n \in \{IP_{j,m}(t_{ScoreIPP}) \cap IP_{k,m}(t_{ScoreIPP})\}} \Delta Pos_{j,k,m,n}(t_{ScoreIPP})^2\right]\right\}}$$

$$\Delta Time_{RMS\_j,k}(t_{ScoreIPP}) = \sqrt{\frac{1}{N}\sum_{m=1}^{N}\left\{\frac{1}{f_N\left[IP_{j,m}(t_{ScoreIPP}) \cap IP_{k,m}(t_{ScoreIPP})\right]}\left[\sum_{n \in \{IP_{j,m}(t_{ScoreIPP}) \cap IP_{k,m}(t_{ScoreIPP})\}} \Delta Time_{j,k,m,n}(t_{ScoreIPP})^2\right]\right\}}$$

For all participating unit pairs, all true launch points covered in common by the pairs, all runs, and all scoring times, plot the statistical density and cumulative distribution of the $\Delta Pos_{j,k,m,n}(t_{ScoreIPP})$ and $\Delta Time_{j,k,m,n}(t_{ScoreIPP})$ statistics.

Higher-level roll-ups of this MOP may be computed as required by the user.

Pairwise Cross-Platform Commonality History: Booster Type Estimate Differences

Definition: TBD(For pairs of participating units, the norm of the difference vector between boost type probability vectors, divided by the sum of the norms of the boost type probability vectors.)

Calculation: For each pair j,k of participating units, use the $LP_{j,m}(t_{ScoreLPE})$ and $LP_{k,m}(t_{ScoreLPE})$ sets determined during the calculation of the ratio of non-common launch points.

For each schedule LPE scoring time $t_{ScoreLPE}$, each run m, each pair of participating units, and each true ballistic missile launch point n in the set $LP_{j,m}(t_{ScoreLPE}) \cap LP_{k,m}(t_{ScoreLPE})$, record the following

$$\Delta BC_{j,k,m,n}(t_{ScoreLPE}) = \frac{\| \mathbf{BT}_{j,m,n}(t_{ScoreLPE}) - \mathbf{BT}_{k,m,n}(t_{ScoreLPE}) \|^2}{\| \mathbf{BT}_{j,m,n}(t_{ScoreLPE}) \|^2 + \| \mathbf{BT}_{k,m,n}(t_{ScoreLPE}) \|^2}$$

where $\mathbf{BT}_{j,m,n}(t_{ScoreLPE})$ and $\mathbf{BT}_{k,m,n}(t_{ScoreLPE})$ are the appropriate booster typing probability vectors associated by participants j and k for the same launch point n on run m. (See the calculation of Track Commonality for the convention for accounting for a “no statement” response.)

At each scheduled scoring time $t_{ScoreLPE}$ for pairs of participating units of interest, compute and plot over the scenario time the average booster classification difference over all true launch points n covered in common and over all runs m, i.e.,

$$\Delta BC_{Aver\_j,k}(t_{ScoreLPE}) = \frac{1}{N}\sum_{m=1}^{N}\left\{\frac{1}{f_N\left[LP_{j,m}(t_{ScoreLPE}) \cap LP_{k,m}(t_{ScoreLPE})\right]}\left[\sum_{n \in \{LP_{j,m}(t_{ScoreLPE}) \cap LP_{k,m}(t_{ScoreLPE})\}} \Delta BC_{j,k,m,n}(t_{ScoreLPE})\right]\right\}$$

For all participating unit pairs, all true launch points covered in common by the pairs, all runs, and all scoring times, plot the statistical density and cumulative distribution of the $\Delta BC_{j,k,m,n}(t_{ScoreLPE})$ statistics.

Higher-level roll-ups of this MOP may be computed as required by the user.

Pairwise Cross-Platform Commonality History: Post-Boost Object Classification Differences

Definition: TBD(For pairs of participating units, the norm of the difference vector between post-boost object classification probability vectors, divided by the sum of the norms of the post-boost object classification probability vectors.)

Calculation: For each pair j,k of participating units, use the $TN_{j,m}(t_{ScoreTrk})$ and $TN_{k,m}(t_{ScoreTrk})$ sets determined for the rate of non-common track numbers.

For each scoring time, each run, each pair of participating units, and each track number n in the set $TN_{j,m}(t_{ScoreTrk}) \cap TN_{k,m}(t_{ScoreTrk})$, record the following

$$\Delta OC_{j,k,m,n}(t_{ScoreTrk}) = \frac{\| \mathbf{OC}_{j,m,n}(t_{ScoreTrk}) - \mathbf{OC}_{k,m,n}(t_{ScoreTrk}) \|^2}{\| \mathbf{OC}_{j,m,n}(t_{ScoreTrk}) \|^2 + \| \mathbf{OC}_{k,m,n}(t_{ScoreTrk}) \|^2}$$

where $\mathbf{OC}_{j,m,n}(t_{ScoreTrk})$ and $\mathbf{OC}_{k,m,n}(t_{ScoreTrk})$ are the appropriate post-boost object classification probability vectors held by participants j and k for the same track number n on run m. (See the calculation of Track Commonality for the convention for accounting for a "no statement" response.)

At each scheduled scoring time $t_{ScoreTrk}$ for pairs of participating units of interest, compute and plot over the scenario time the average post-boost object classification difference over all track numbers n in common and over all runs m, i.e.,

$$\Delta OC_{Aver\_j,k}(t_{ScoreTrk}) = \frac{1}{N}\sum_{m=1}^{N}\left\{\frac{1}{f_N\left[TN_{j,m}(t_{ScoreTrk}) \cap TN_{k,m}(t_{ScoreTrk})\right]}\left[\sum_{n \in \{LP_{j,m}(t_{ScoreTrk}) \cap LP_{k,m}(t_{ScoreTrk})\}} \Delta OC_{j,k,m,n}(t_{ScoreTrk})\right]\right\}$$

For all participating unit pairs, all track numbers in common by the pairs, all runs, and all scoring times, plot the statistical density and cumulative distribution of the $\Delta OC_{j,k,m,n}(t_{ScoreTrk})$ statistics.

Higher-level roll-ups of this MOP may be computed as required by the user.

System Loading: Communications Data Loading

Definition: Sum of data rates input from all platforms into the communications function for distribution to other remote platforms.

Calculation: Define the messages and their bit counts for each type of measurement, track report, LPE report, IPP report, or other generated data that are distributed within the system. Initialize $TC_n(t_{ScoreTrk}) = 0$.

For run n and scenario run time interval $t_{ScoreTrk\_last} < t < t_{ScoreTrk}$, each time a message is "sent" from any platform for distribution to other remote nodes, increment $TC_n(t_{ScoreTrk})$ by the bit count of the message.

Average over all runs as $TC(t_{ScoreTrk}) = \frac{1}{N}\sum_{n=1}^{N} TC_n(t_{ScoreTrk})$

Compute and plot the interval input data rates (versus time into the scenario) as $DR(t_{ScoreTrk}) = \frac{TC(t_{ScoreTrk})}{t_{ScoreTrk} - t_{ScoreTrk\_last}}$, and highlight the peak interval and rate.

Higher-level roll-ups of this MOP may be computed as required by the user.

APPENDIX B

SIAP Metrics and TAMD CRD Requirements

This appendix documents, for future use, TAMD CRD requirements pertaining to the SIAP and to TBM EW, and correspondences between these requirements and the ballistic missile SIAP metrics defined in this report.

It should be noted that the SIAP SE is attempting to define ballistic missile metrics that can be applied to Integrated Missile Defense (IMD). Thus, the scope of the present effort extends beyond that of the TAMD CRD.

Tables 1, 2, and 3 show the TAMD CRD KPP Attributes and Metrics and their associated requirements mapped to SIAP attributes and metrics from this technical report (and from SIAP SE TF Technical Report 2001-001 as applicable).

The tables should not be interpreted to imply that there is a one-to-one relationship of the TAMD CRD requirements and the metrics defined in this technical report. The mapping shows that there is enough commonality between some of the requirements and the SIAP metrics to suggest that one could be derived from the other or that minor modifications to the SIAP metrics could support the TAMD CRD requirements. Because the language of the TAMD CRD is at points difficult to relate to the SIAP SE's quantitative assessment framework, it may be necessary to obtain further guidance from USJFCOM on the intent of certain CRD requirements before this mapping can be finalized. It is expected that this goal will be accomplished by the publication of Version 2.0 of this technical report, scheduled for late in 2002.

The tables also show that most, but not all, of the TAMD CRD Attributes and Metrics are addressed within the Ballistic Missile SIAP Metrics Technical Report.

The TAMD CRD KPPs and Attributes

Tables 4 – 24 provide descriptions of each of the TAMD CRD SIAP KPPs, TBM Early Warning KPPs, and TBM Early Warning Attributes and Metrics. Each table provides a definition of the topic, supporting definitions of terms used, and a discussion of the topic relevant to ballistic missile metrics.

The following points derived from the warfighter perspective provided by the TAMD CRD are noted for future consideration:

1. Neither Area of Interest nor Area of Influence may be the most appropriate term for describing the aerospace volume in which BMD metrics should be calculated. Area of Responsibility as defined in Joint Pub 1-02 and Joint Pub 3-0 should be considered. The discussions on TBM Attack Early Warning in the TAMD CRD explicitly states that the aerospace volume of concern includes the area of responsibility.

2. The definition of significant damage currently used in the Technical Report may be too narrow. An alternative definition is suggested in Table 9.
3. The definition of objects of interest needs careful consideration. From the warfighter perspective, all objects capable of significant damage may be the appropriate selection. However, from the perspective of the defense system operator, this may not be adequate. The selection of objects and tracks for scoring may be different for CRD satisfaction than for weapon system performance.
4. The intent of the TAMD CRD with respect to Early Warning is to provide warning in terms of launch alerts and impact prediction to the applicable affected forces. This implies that the relevant metrics need to be calculated at the end user. This applies as well to launch time estimates and impact time estimates. Thus, the SIAP SE needs to define metrics that account for message dissemination to the end user for these cases.
5. Similarly, overwarning is measured as spurious alerts to the applicable affected forces. Algorithmic determination that an impact point prediction is invalid should not eliminate that prediction if it has been disseminated to the applicable forces. Any impact point message that is sent to the applicable forces is a message that they must deal with.
6. The requirement for impact time accuracy is one-sided. That is, at the specified time before actual impact, the difference between the predicted impact time and the actual impact time must be a number greater than or equal to zero.

TAMD CRD Requirements Mapping to SIAP SE Metrics

Table 1: TAMD CRD SIAP KPP Attributes and Metrics

	Attribute	Associated Requirements	Relevant SIAP SE Metric (Appendix A)
1	Completeness	Percent detected and tracked Available and exploitable to JFC	Object track completeness *
2	Ambiguity	Average tracks per object Percent tracks represent truth objects	Track ambiguity Spurious track mean ratio
3	Continuity	Track time without any drops, duals/splits, merges, or swaps	Cumulative switches of tracks on distinct reportable objects Cumulative switches of tracks on reportable object threads Cumulative broken tracks on reportable objects Cumulative broken tracks on reportable object threads Normalized longest duration valid track segment on distinct reportable objects Normalized longest duration valid track segment on distinct reportable object threads

Table 2: TAMD CRD TBM Attack Early Warning KPP

	Event	Associated Requirements	Relevant SIAP SE Metric (Appendix A)
1	Launch report	Time from first detection Accuracy Accuracy confidence	LPE delay LPE position and time accuracy LPE position and time uncertainty consistency
2	Initial IPT & IPP	Time after BO	IPP warning time
3	Update IPT & IPP report	Time available before impact Accuracy at this time Accuracy confidence at this time	IPP warning time IPP position and time accuracy IPP position and time uncertainty consistency

Table 3: TAMD CRD TBM Attack Early Warning Attributes and Metrics

	Event	Associated Requirements	Relevant SIAP SE Metric (Appendix A)
1	Detect	Probability of detection	Launch point compeletness
2	Initial Boost Phase Report	Time after cloud break Time after launch Accuracy of LPE Accuracy of LTE Probability of correct typing Probability of warning Probability of overwarning/false alarm Probability of receipt	LPE delay LPE delay LPE position accuracy LPE time accuracy Booster typing correctness IPP completeness ** **
3	Boost Phase Update	Time after previous report	***
4	Final Boost Phase	Time after burnout	***
5	Initial Post-Boost Report	Time after burnout IPP accuracy	Booster burnout estimate delay
6	Post-Boost Phase Update Reports	Time after previous report IPP accuracy before impact IPT accuracy before impact	***
7	Final Post-Boost Phase Update Report	Time after track loss Probability of receipt	***

* - to be addressed in 2002 update of SIAP SE TF Technical Report 2001-001

** - expected to be derivable from Ballistic Missile SIAP attribute measures; details omitted pending USJFCOM clarification of CRD requirements

*** - SIAP metrics do not directly assess update times, but the latter are extractable from data recorded in the scoring of other metrics. Additional SIAP MOPs will be introduced as needed to accommodate any USJFCOM clarification of requirements.

SIAP KPPs

Table 4. TAMD CRD SIAP KPP - Completeness

Completeness		
Definition The measure of the portion of ground truth tracks that are included in the SIAP. The metrics that describe this attribute are: percent of threat objects in track upon entering the area of influence; percent of total objects in track upon entering the area of influence; and percent of primary/secondary systems as defined by the scenario, to which tracks are available and exploitable.		TAMD CRD
Measure Percent of ground truth threat objects detected and tracked upon entering the area of influence Percent of ground truth aircraft detected and tracked upon entering the area of influence Available and exploitable by percent of the primary/secondary systems available to the JFC		
Supporting Definitions		
Ground truth	100 percent of all entities in the given battlespace at their true location. This does not imply surveillance sensors have detected the object.	TAMD CRD Glossary
Detection	1. In tactical operations, the perception of an object of possible military interest but unconfirmed by recognitin. 2. In surveillance, the determination and transmission by a surveillance system that an event has occurred.	From Joint Pub 1-02
Track	A series of related contacts that can be displayed to support planning and tactical actions.	Derived from Joint Pub 1-02 and TAMD CRD
Area of Influence	A geographical area wherein a commander is directly capable of influencing operations by maneuver or fire support systems normally under the commander's command or control.	Joint Pub 1-02
Area of Interest	That area of concern to the commander, including the area of influence, areas adjacent thereto, and extending into enemy territory to the objectives of current or planned operations. This area also includes areas occupied by enemy forces who could jeopardize the accomplishment of the mission.	Joint Pub 1-02
Area of Responsibility	The geographical area associated with a combatant command within which a combatant commander has authority to plan and conduct operations. Also called AOR.	Joint Pub 1-02 and Joint Pub 3-0
Ballistic Missile	Any missile which does not rely upon aerodynamic surfaces to produce lift and consequently follows a ballistic trajectory when thrust is terminiated.	Joint Pub 1-02
Theater Missile	A theater missile, with a range of 3500km or less, which may be a ballistic missile, a cruise missile, or an air-to-surface missile (not including short-range, non-nuclear, direct fire missiles, bombs, or rockets such as Maverick or wire-guided missiles), whose target is within a theater of operations.	Joint Pub 1-02 as modified by ABM Treaty language
Air-breathing Threat:	For purposes of this document, ABTs include CMs, fixed- and rotary-wing aircraft, ASMs, LCRs and UAVs.	TAMD CRD Glossary
Significant Damage	Loss of the use of a military facility or capability for a period of time sufficient to disrupt the conduct of operations or the loss of human lives.	Recommended definition
Available	Accessible to the TAMD primary/secondary systems either continuously or on demand.	Recommended definition
Exploitable	Characterized by information adequate to support tactical actions where tactical actions include activation of systems to gather additional information on objects of interest.	Recommended definition
Issues/Discussion Because the desire is to provide attributes and metrics that apply to the full spectrum of ballistic missile defense, it is recommended that the area of influence be interpreted as the area of responsibility from Joint Pub 1-02. Again, because of the interest in addressng the full scope of ballistic missile defense, the ground truth threat objects should include all ballistic objects with the exception of LCRs. The ground truth threat objects should include all ballistic missiles and ballistic objects that are capable of inflicting significant damage upon impact. This is in contrast with the completeness requirement for air-breathers where the goal is to track every object.		

Table 5. TAMD CRD SIAP KPP - Ambiguity

Ambiguity		
Definition		
The measure of the clarity of tracks in the SIAP. The metrics which describe this attribute are: average number of tracks/air object in track at any given time; and percent of SIAP tracks that represent distinct ground truth objects.		JTAMDO from TAMD CRD Glossary
Measure		
Average number of tracks per air object Percent of tracks represent distinct ground truth objects		
Supporting Definitions		
Average	Ratio of tracks to ground truth objects within the area of influence at any instant in	Recommended definition
Issues/Discussion		
For BMD metrics, the average number of tracks per air object is the ratio of tracks on the objects of interest to the total number of these objects. The objects of interest include boosters, deployment vehicles, warheads, and objects capable of inflicting significant damage upon impact. This quantity is to be measured at points in time. The percent of tracks that represent distinct ground truth objects is the ratio of tracks on the objects of interest to the sum of those tracks and all other tracks identified as being tracks on the objects of interest.		

Table 6. TAMD CRD SIAP KPP - Continuity

Continuity		
Definition		
The measure of how well the SIAP maintains tracks, track numbers, tracks identification, and other track attributes over time. The metrics which describe this attribute are: time without any drops, duals/splits, and merges; and time without any track number or identification swapping.		JTAMDO from TAMD CRD Glossary
Measure		
Track time without any track drops, duals/splits, merges, or swaps		
Supporting Definitions		
Drop	Cessation of a track on a ground truth object while that object continues to exist within the area of influence.	Recommended definition
Dual	Presence of more than one reported track on a ground truth object within the Single Integrated Air Picture.	Recommended definition
Merge	The formation of a single track from multiple tracks when those tracks were initially associated with one or more ground truth objects.	Recommended definition
Split	Generation of multiple tracks on a ground truth object that initially was represented by a single track.	Recommended definition
Swap	An exchange of either the track numbers or amplifying data on tracks associated with distinct ground truth objects.	Recommended definition
Issues/Discussion		
For BMD metrics, continuity of tracks on the objects of interest is the focus.		

TBM Early Warning KPPs

Table 7. TAMD CRD SIAP TBM Early Warning KPP - TBM Launch Detection

Initial TBM Launch Report		
Definition		
A message containing, at a minimum, the following information: 1) estimated time of launch, 2) estimated point of launch and, 3) launch azimuth.		Recommended definition
Measure		
Time from first detection Azimuth accuracy of missile heading Confidence level in heading accuracy		
Supporting Definitions		
Detection	1. In tactical operations, the perception of an object of possible military interest but unconfirmed by recognitin. 2. In surveillance, the determination and transmission by a surveillance system that an event has occurred.	From Joint Pub 1-02
Launch Azimuth	Missile launch direction measured in degrees clockwise from true north.	As revised from BMDO Glossary
Issues/Discussion		
The TAMD CRD Interoperability KPP requires that the early warining be delivered to the applicable affected forces. Thus, the time is to be measured to the delivery of the launch report to the information user.		

Table 8. TAMD CRD SIAP TBM Early Warning KPP - TBM Impact Time and Point

Impact Time and Impact Point Reports		
Definition		
A message containing, at a minimum, the following information: 1) estimated time of impact, and 2) estimated point of impact.		Recommended definition
Measure		
Seconds after booster burn-out for initial report Seconds before impact for final report Predicted impact point accuracy for final report Confidence level in impact point for final report		
Supporting Definitions		
Impact Point Accuracy	The distance from actual missile impact to predicted missile impact point. Methodologies for establishing impact point accuracy may take a number of forms: An ellipse with uncertainties along the semi-major and semi-minor axes, an azimuth "wedge" subtended by maximum and minimum predicted range limits, a geometric figure that contains impact point predictions and associated errors from all early warning elements.	USSPACECOM
Impact Point Prediction	Prediction of the point on the earth's surface where a specific RV will impact, usually specified in terms of the circular error probable. The estimate includes the perturbing effects of the atmosphere and resultant uncertainties.	BMDO Glossary
Impact Point Report Time	A measure of the timeliness or latency of reporting impact point information to end-users. Two methods lend themselves to reporting this information: counting up from computed time of launch, and counting down to predicted time of impact.	USSPACECOM
Issues/Discussion		
The TAMD CRD Interoperability KPP requires that the early warining be delivered to the applicable affected forces. Thus, the time is to be measured to the delivery of the impact point report to the information user.		

Early Warning Attributes

Table 9. TAMD CRD Early Warning Attributes - Probability of Detection

Probability of Detection (single event)		
Definition		
The probability of initial indication by any one of a variety of sensors that a booster has been launched from some point on the surface of the earth, with initial characterization of the booster type.		Recommended definition
Measure		
Fraction		
Supporting Definitions		
Detection	1. In tactical operations, the perception of an object of possible military interest but unconfirmed by recognition. 2. In surveillance, the determination and transmission by a surveillance system that an event has occurred.	From Joint Pub 1-02
Launch Detection	Launch Detection: Initial indication and characterization by any one of a variety of sensors that a theater missile has been launched.	TAMD CRD Glossary
Launch Detection	Initial indication by any one of a variety of sensors that a booster has been launched from some point on the surface of the earth, with initial characterization of the booster type.	BMDO Glossary
Issues/Discussion		
A ballistic missile should be considered detected when a surveillance system has transmitted a launch report.		

Table 10. TAMD CRD Early Warning Attributes - Probability of Warning

Probability of Warning (given detection)		
Definition		
Probability that a message providing information on the estimated time to impact and estimated impact point and/or geographical area at risk is delivered to the applicable affected forces.		Recommended definition
Measure		
Fraction		
Supporting Definitions		
Warning	Warning consists of information on the estimated time to impact and estimated impact point and/or geographical area at risk.	TAMD CRD
Impact Point Prediction	Prediction of the point on the earth's surface where a specific RV will impact, usually specified in terms of the circular error probable. The estimate includes the perturbing effects of the atmosphere and resultant uncertainties. (BMDO Glossary)	TAMD CRD
Issues/Discussion		
The TAMD CRD Interoperability KPP requires that the early warning be delivered to the applicable affected forces. Thus, warning is accomplished when the warning message is delivered to the applicable affected forces.		

Table 11. TAMD CRD Early Warning Attributes - Probability of Overwarning/False Alarm

Probability of Overwarning/False Alarm		
Definition		
Probability that a message providing information on the estimated time to impact and estimated impact point and/or geographical area at risk is incorrectly delivered to the applicable affected forces.		Recommended definition
Measure		
Fraction		
Supporting Definitions		
Warning	Warning consists of information on the estimated time to impact and estimated impact point and/or geographical area at risk.	TAMD CRD
Probability of false alarm	The probability that a missile launch event will be reported when no event has occurred. An equitable balance must be maintained between assured warning and false alarm rate. Further, once a report is known to be false, an immediate retraction of this warning is required. (USSPACECOM)	TAMD CRD
Issues/Discussion		
The TAMD CRD Interoperability KPP requires that the early warning be delivered to the applicable affected forces. Thus, warning is accomplished when the warning message is delivered to the applicable affected forces.		

Table 12. TAMD CRD Early Warning Attributes - Probability of Correct Typing

Probability of Correct Typing-Prior to First Report		
Definition		
Probability that a ballistic missile is correctly identified for purposes of establishing impact point metrics.		Recommended definition
Measure		
Fraction		
Supporting Definitions		
Type	That which possesses of exemplifies characteristic qualities.	CID CRD Glossary
Probablity of Correct Typing	Current and projected systems that process space-based data rely on typing a missile correctly and comparing it to a known profile in order to provide accurate TBM attack early warning information, especially in-flight and impact point metrics. (USSPACECOM)	TAMD CRD
Issues/Discussion		
In the context of early warning, typing should support the impact point metrics including the geographical area at risk.		

Table 13. TAMD CRD Early Warning Attributes - Launch Point Accuracy

Launch Point Accuracy-Initial Boost Phase Update Report		
Definition		
Radial distance from the actual launch point to the estimated launch point within the specified confidence bound.		Recommended definition
Measure		
Distance within confidence bound		
Supporting Definitions		
Launch Point Report	A message containing, at a minimum, the following information: 1) estimated time of launch, 2) estimated point of launch and, 3) launch azimuth.	Recommended definition
Launch Point Accuracy	The radial distance from actual launch point to predicted launch point, given as circular error probable (CEP). (USSPACECOM)	TAMD CRD
Launch Point Probability of Inclusion	The probability that the actual launch point falls within the predicted launch point CEP. (USSPACECOM)	TAMD CRD
Launch Azimuth	Missile launch direction measured in degrees clockwise from true north. (As revised from BMDO Glossary)	TAMD CRD
Issues/Discussion		
None.		

Table 14. TAMD CRD Early Warning Attributes - Launch Time Accuracy

Launch Time Accuracy-Initial Boost Phase Update Report		
Definition		
Time from the actual launch time to the estimated launch time within the specified confidence bound.		Recommended definition
Measure		
Time		
Supporting Definitions		
Launch Point Report	A message containing, at a minimum, the following information: 1) estimated time of launch, 2) estimated point of launch and, 3) launch azimuth.	Recommended definition
Launch Time Accuracy	The difference between actual and predicted time of launch.	USSPACECOM
Issues/Discussion		
None.		

Table 15. TAMD CRD Early Warning Attributes - Impact Time Accuracy

Impact Time Accuracy at Specified Time Before Impact		
Definition		
Difference between the actual impact time and the predicted impact time.		Recommended definition
Measure		
Time		
Supporting Definitions		
Impact Point Report Time	A measure of the timeliness or latency of reporting impact point information to end-users. Two methods lend themselves to reporting this information: counting up from computed time of launch, an counting down to predicted time of impact.	USSPACECOM
Issues/Discussion		
The accuracy requirement for predicted impact time is one-sided. That is, at the specified time before actual impact, the difference between the predicted impact time and the actual impact time must be a number between zero and the accuracy requirement.		

Table 16. TAMD CRD Early Warning Attributes - Impact Point Accuracy After Burnout

Impact Point Accuracy at a Specified Time after Burn-out		
Definition		
Radial distance from the actual impact point to the predicted impact point within the specified confidence bound.		Recommended definition
Measure		
Distance within confidence bound		
Supporting Definitions		
Impact Point Prediction	Prediction of the point on the earth's surface where a specific RV will impact, usually specified in terms of the circular error probable. The estimate includes the perturbing effects of the atmosphere and resultant uncertainties.	BMDO Glossary
Impact Point Accuracy	The distance from actual missile impact to predicted missile impact point. Methodologies for establishing impact point accuracy may take a number of forms: An ellipse with uncertainties along the semi-major and semi-minor axes, an aximuth "wedge" subtended by maximum and minimum predicted range limits, a geometric figure that contains impact point predictions and associated errors from all early warning elements.	USSPACECOM
Impact Point Report	"Other desired characteristics of EW accuracy are warhead typing notification/situational awareness, warhead effects (including debris) prediction to affected units, precise warning to affected units versus general warning to the entire theater/sector, and rapid 'all-clear' message issuance."	JTAMD White Paper on Early Warning for Passive Defense, 30 March 1999
Issues/Discussion		
None.		

Table 17. TAMD CRD Early Warning Attributes - Impact Point Accuracy at Time Before Impact

Impact Point Accuracy at Specified Time Before Impact		
Definition		
Radial distance from the actual impact point to the predicted impact point within the specified confidence bound.		Recommended definition
Measure		
Distance within confidence bound		
Supporting Definitions		
Impact Point Prediction	Prediction of the point on the earth's surface where a specific RV will impact, usually specified in terms of the circular error probable. The estimate includes the perturbing effects of the atmosphere and resultant uncertainties.	BMDO Glossary
Impact Point Accuracy	The distance from actual missile impact to predicted missile impact point. Methodologies for establishing impact point accuracy may take a number of forms: An ellipse with uncertainties along the semi-major and semi-minor axes, an aximuth "wedge" subtended by maximum and minimum predicted range limits, a geometric figure that contains impact point predictions and associated errors from all early warning elements.	USSPACECOM
Impact Point Report	"Other desired characteristics of EW accuracy are warhead typing notification/situational awareness, warhead effects (including debris) prediction to affected units, precise warning to affected units versus general warning to the entire theater/sector, and rapid 'all-clear' message issuance."	JTAMD White Paper on Early Warning for Passive Defense, 30 March 1999
Issues/Discussion		
None.		

Table 18. TAMD CRD Early Warning Attributes - Reporting Times for Initial Boost Phase Report

Reporting Times-Initial Boost Phase Report		
Definition		
For threshold performance, the time after cloud-free line of sight that the boost phase report is disseminated. For objective performance, the time after launch that the boost phase report is disseminated.		Recommended definition
Measure		
Time		
Supporting Definitions		
Boost Phase Report	A message containing, at a minimum, the following information: 1) estimated time of launch, 2) estimated point of launch, and 3) type of missle launched.	Recommended definition
Issues/Discussion		
None.		

Table 19. TAMD CRD Early Warning Attributes - Reporting Times for Boost Phase Report Updates

Reporting Times-Boost Phase Update Reports		
Definition		
Time between boost phase reports		Recommended definition
Measure		
Time		
Supporting Definitions		
Boost Phase Report	A message containing, at a minimum, the following information: 1) estimated time of launch, 2) estimated point of launch, and 3) type of missle launched.	Recommended definition
Issues/Discussion		
None.		

Table 20. TAMD CRD Early Warning Attributes - Reporting Times for Final Boost Phase Report

Reporting Times-Final Boost Phase Report		
Definition		
Time after booster burn-out that the last boost phase report is sent.		Recommended definition
Measure		
Time		
Supporting Definitions		
Boost Phase Report	A message containing, at a minimum, the following information: 1) estimated time of launch, 2) estimated point of launch, and 3) type of missle launched.	Recommended definition
Issues/Discussion		
None.		

Table 21. TAMD CRD Early Warning Attributes - Reporting Times for Initial Post Boost Phase Report

Reporting Times-Initial Post Boost Phase Report		
Definition		
Time after booster burn-out that the first post boost phase report is sent.		Recommended definition
Measure		
Time		
Supporting Definitions		
Post Boost Report	A message containing, at a minimum, the following information: 1) estimated time of impact, 2) estimated point of impact, 3) object type.	Recommended definition
	"Other desired characteristics of EW accuracy are warhead typing notification/situational awareness, warhead effects (including debris) prediction to affected units, precise warning to affected units versus general warning to the entire theater/sector, and rapid 'all-clear' message issuance."	JTAMD White Paper on Early Warning for Passive Defense, 30 March 1999
Issues/Discussion		
None.		

Table 22. TAMD CRD Early Warning Attributes - Reporting Times for Post Boost Phase Update Reports

Reporting Times-Post Boost Phase Update Reports		
Definition		
Time between post boost phase reports		Recommended definition
Measure		
Time		
Supporting Definitions		
Post Boost Report	A message containing, at a minimum, the following information: 1) estimated time of impact, 2) estimated point of impact, 3) object type.	Recommended definition
	"Other desired characteristics of EW accuracy are warhead typing notification/situational awareness, warhead effects (including debris) prediction to affected units, precise warning to affected units versus general warning to the entire theater/sector, and rapid 'all-clear' message issuance."	JTAMD White Paper on Early Warning for Passive Defense, 30 March 1999
Issues/Discussion		
None.		

Table 23. TAMD CRD Early Warning Attributes - Reporting Times for Final Post Boost Phase Reports

Reporting Times-Final Post Boost Phase Report		
Definition		
Time after loss of track on a ballistic object capable of inflicting significant damage upon impact.		Recommended definition
Measure		
Time		
Supporting Definitions		
Post Boost Report	A message containing, at a minimum, the following information: 1) estimated time of impact, 2) estimated point of impact, 3) object type.	Recommended definition
	"Other desired characteristics of EW accuracy are warhead typing notification/situational awareness, warhead effects (including debris) prediction to affected units, precise warning to affected units versus general warning to the entire theater/sector, and rapid 'all-clear' message issuance."	JTAMD White Paper on Early Warning for Passive Defense, 30 March 1999
Issues/Discussion		
None.		

Table 24. TAMD CRD Early Warning Attributes - Probability of Initial and Final Message Receipt

Probability of Initial & Final Message Receipt		
Definition		
Probability that the initial boost phase report and the final post boost report are received at the applicable affected forces.		Recommended definition
Measure		
Fraction		
Supporting Definitions		
Boost Phase Report	A message containing, at a minimum, the following information: 1) estimated time of launch, 2) estimated point of launch, and 3) type of missle launched.	Recommended definition
Post Boost Report	A message containing, at a minimum, the following information: 1) estimated time of impact, 2) estimated point of impact, 3) object type.	Recommended definition
	"Other desired characteristics of EW accuracy are warhead typing notification/situational awareness, warhead effects (including debris) prediction to affected units, precise warning to affected units versus general warning to the entire theater/sector, and rapid 'all-clear' message issuance."	JTAMD White Paper on Early Warning for Passive Defense, 30 March 1999
Issues/Discussion		
None.		

Miscellanous Notes and Definitions

Early Warning applies to any ballistic threat launched from, over-flying, or projected to impact in the designated area of responsibility, JOA, and/or area of interest.

Launch Azimuth: Missile launch direction measured in degrees clockwisw from the true north. (As revised from BMDO Glossary).

Launch Point Accuracy: The radial distance from actual launch point to predicted launch point, given as circular error probable (CEP). (USSPACECOM)